# Referential choices in language production

Vogels, J.

*Publication date:*
2014

*Document Version*
Publisher's PDF, also known as Version of record

# Referential choices in language production

The role of accessibility

Jorrig Vogels

Referential choices in language production: The role of accessibility

Jorrig Vogels
Ph.D. thesis
Tilburg University

TiCC Ph.D. series no. 33

ISBN: 978-94-6203-560-7
Print: CPI Wöhrmann print service
Cover design: Milan Vogels

# Referential choices in language production

## The role of accessibility

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan Tilburg University,
op gezag van de rector magnificus,
prof. dr. Ph. Eijlander,
in het openbaar te verdedigen ten overstaan van een
door het college voor promoties aangewezen commissie
in de aula van de Universiteit
op woensdag 23 april 2014 om 14.15 uur

door

**Jorrig Vogels**

geboren op 13 januari 1986 te Nijmegen

**Promotores:**

Prof. Dr. E. J. Krahmer

Prof. Dr. A. A. Maes

**Overige leden van de Promotiecommissie:**

Dr. A. Gatt

Dr. R. P. G. van Gompel

Prof. Dr. P. Hendriks

Dr. E. M. Kaiser

Prof. Dr. L. G. M. Noordman

*[I]n a conceptualist theory, reference is taken to be at its foundation dependent on a language user—just as relativistic physics takes distances and times to be dependent on an observer's inertial frame.* – Ray Jackendoff (2002: 304)

*Je moet gewoon niet te diep nadenken, en dan klopt alles.* – Herman Finkers (Geen Spatader Veranderd, 1995)

# Acknowledgments

When I finished my Master's thesis in Linguistics at the Radboud University Nijmegen in September 2009, it had taken me one year longer than was scheduled in the curriculum. My second reader, Helen de Hoop, commented that if it were to happen that I would need an extra year for my Ph.D. as well, she would be willing to speak in favor of me, because she knew it would be worth it. This appeared not to be necessary: Even to my own surprise, I finished my dissertation in just slightly over four years, and I am not unhappy with the result. Perhaps spending another year on it would have made it even better, but to quote my high school teacher of classical languages: "You can get from a 6 to an 8 within a reasonable amount of time and effort, but getting from an 8 to a 10 requires a lot more." In any case, thank you Helen for your faith in my capacities.

Two important reasons why I successfully completed my Ph.D. in time are called Fons Maes and Emiel Krahmer. While they are both experienced Ph.D. supervisors, I do not think they have supervised many theses together. I would advise them to do this more often, because I experienced it as a very fruitful combination. With Fons, I could have interesting theoretical discussions on possible explanations for a strange effect in my data, Fons occasionally pointing to relevant research he had done twenty years ago. When at the end of such a discussion I was convinced I had to do a new experiment with 32 independent variables, Emiel quickly brought me back to the real world and helped me setting up a study that was actually feasible. I thank my supervisors for their great support, for their kindness, and their quick and helpful responses. I remember that one time I had just finished a paper and sent it to my supervisors with the idea of spending a few days doing other stuff. Unfortunately, the next morning their comments already entered my mailbox. It makes you think that professors have a hidden drawer somewhere from which they can pick up some extra time when needed.

When I entered the DCI department at Tilburg University in 2009, I immediately felt at home. I liked the casual, informal atmosphere and the friendly people. I enjoyed the cookie- as well as the PhD-meetings, the PhDinners and the visits to Malle. I even joined the Malle band, which has now been successful for four years in a row. I shared

an office with fellow-freshman Constantijn Kaland, who is not only an expert in phonetics, but also has a quite extensive repertoire of funny noises and accents. Our four years together can be characterized by two short conversations: "Hoe ver ben jij al met je schroefpift? O, ik moet alleen nog een paar stukjes tikken" and "Vind je het goed als ik een raam openzet? Ja, dat raam bijvoorbeeld"[1]. In our final year, we were assigned a mystery officemate, which turned out to be Mariana from Portugal. But due to a little accident, we suddenly had a fourth person in our office as well. This made it time for me to leave and join Phoebe and Sylvia. I would like to thank all five officemates for the nonsense as well as for the more serious discussions. I hope my future officemates, if any, will be as fun.

Sometimes I also walked out of my office to see other people, or other people walked into mine to see me. Some of them I want to mention in particular: With a number of people I could share my interest in reference and language production: Adriana, Hans, Ingrid, Jette, Marieke, Martijn G. and Ruud K., thank you for all the fruitful meetings and interesting discussions. For more general discussions on language, linguistics, and methodology, as well as for general silliness, I would like to thank Lisette (I hope Mol & Vogels (20??) will become reality one day), Lisanne (still sorry I compared you to a freight train), Naomi (thanks for sharing your formal semantics library), Yan (thank you for your delicious Chinese cooking), and the other 4th floor Ph.D. students: Alain, Emmelyn, Karin, Lieke, Mandy, Rick and Ruud M. You have all been great colleagues.

For two people I did not even have to leave my office to communicate with them: Thank you Maria and Véronique for the shouting across the corridor and the casual as well as the more profound conversations. And Véronique, thank you for wanting to be one of my paranymphs. A little further from shouting distance were Carel, Kiek, and Marc: Thank you for being the gatekeepers of our part of the corridor. I would also like to thank Jacintha and Lauraine for their support and cheerfulness, and Jacqueline and Rein for their help with everything lab-related. Finally, my four years at DCI were also very musical. A warm thank you to Anja and Anne, and to the rest of the Malle band: Juliette, Leonoor, Mandy, Marije, Martijn B., Menno and Ruud K. Our musical rendezvouses were a pleasure, and I hope we can continue them in the future.

With some colleagues, I already spoiled it in my first year, however, by locking them up in a room and making pictures of them in embarrassing positions. For that I

[1] Our office had only one window that could be opened.

sincerely apologize to Mandy (a.k.a. 'the blonde girl with the big earrings'), Ruud K., Lisette, Constantijn, Marieke, Rein, Kitty, Marjolijn, Martijn B. and Hans. On the other hand, some of these people are now world-famous, featuring in this dissertation as well as in several other publications. I should also thank university photographer Ben Bergmans here, who made a second series of beautiful pictures (to the excitement of the aforementioned colleagues). For comparison: Figure 2.1. in this dissertation is my own work; Figure 2.4. is Ben's. You may judge for yourself. In addition, I would like to thank Hanneke Schoormans for being the voice-over accompanying the pictures, and Ed Boschman for making these recordings sound crystal clear. I would also like to thank former student-assistants Kristel Bartels and Madelène Munnik for their help with some of my other experiments.

Although I have come to like Tilburg a lot, I am also happy that my world is a bit larger than that. I especially want to thank Geertje van Bergen: After being a great Master's thesis supervisor (for one thing, you taught me how to work with R), you were also a great co-author, and I am proud of our paper; Monique Lamers and Suzan Verberne, I enjoyed working together with you, and I still hope our joint paper will be accepted one day; and Jacolien van Rij, thank you for discussing pronouns with me and for our nice tours across Manhattan. For the study described in Chapter 4, I went to the Meertens Institute in Amsterdam to recruit participants. I would like to thank all participants for volunteering, as well as Ben Hermans, Marc van Oostendorp and Anke van Reenen for facilitating the experiment. For the Flemish part of this study, Fons made recordings of his own family and friends. I'm grateful to these respondents as well, and I hope they are still Fons's friends. I take all responsibility.

This brings me to my own family and friends, without the support of whom this dissertation obviously would not have been as good. Mieke and Léon, thank you for your engagement in what I do and for our discussions about statistics and career prospectives (at the plantsoenendienst). Milan, thank you for having been willing to make a nice cover design even when you did not have the time. Floor, it seems that you were studying in Tilburg too. Ah well, there will be plenty of occasions to meet up in the future. And then there are all the other family members that I bombarded with puzzling pretests. I apologize, and you can now read this book to see what is was all good for.

As for my friends, I would like to mention Leon: Thank you for sharing your ideas about my work, for our scientific, philosophical, personal and rather silly discussions,

for the Herman Finkers quotes and for our cycling and hiking adventures; Daniël, thank you for our trips to the stripbeurs and for being a constant factor that I can always rely on; Marco, thank you for discussing the state of the world with me and for our musical soirées; Sabrina, thank you for wanting to be my other paranymph, for dropping by occasionally in my office for a chat, and for the numerous interesting conversations; Frank, thank you for using your photographer's eye to pick out the photo that as you can see has now made it to the cover; Marlies, Dieter, Maya, Tineke, Noortje, Petra, and everyone I forget, thank you for distracting me from my work when necessary, and for being good friends.

The final words here I would like to address to my *duktig flicka* Josefin. Not only did you greatly improve the text of this dissertation by your 425 detailed and critical comments, you have also been my greatest support since the day we met. Jag tycker om dig så mycket. Du är fantastisk!

# Contents

# Chapter 1

# Introduction

### 1.1. Referential choices

Reference is an essential part of language. When we speak, we talk *about things* (e.g., objects, other people). The act of referring can be seen as forming a link between the speaker's mind and the outside world. For example, a speaker asking 'could you hand me that stapler?' is expressing her[1] intention to get hold of a physical object in the world by referring to that object with a linguistic expression (in this case, the definite noun phrase 'the stapler'). The things we refer to are, however, not always physical objects (including people), nor do they need to be part of the outside world. For example, we can refer to objects that are not present in the direct physical environment ('I left the stapler in the office'), or objects that only exist in our imagination ('the stapler I dreamt about last night'). We can refer to objects that existed in the past ('the cake that I ate yesterday'), or will exist in the future ('the cake that I will bake tomorrow'). We can also refer to events ('last night's dinner party'), locations ('the picturesque town of Tilburg'), and abstract concepts ('the financial crisis'), to name a few. In none of these situations is the thing that is being referred to (the referent) an object in the directly perceivable world. It would therefore be better to say that we refer to *conceptualizations* in our minds, rather than to objects in the outside world (e.g., Jackendoff, 2002; Johnson-Laird, 1983). Even in those cases where the referent *is* present in the world, reference is still mediated by a conceptualization of the referent (which may be wrong, as in 'Could you hand me that stapler?' 'That's not a stapler, that's a hole punch.').

This dissertation is concerned with the process of putting these conceptualizations into language. Although people can refer to concepts denoting all kinds of things, as noted above, this dissertation is confined to reference to concrete entities. In addition, it presents research on language production rather than on comprehension. The reason is that reference production has received less attention than reference resolution in psycholinguistic research, while there is growing evidence that the production and interpretation of referring expressions might not be determined by the same factors (e.g., Kehler, Kertz, Rohde, & Elman, 2008; but cf. also Pickering & Garrod, 2013).

In Levelt's model of language production (Levelt, 1989), a speaker who wants to communicate about a certain entity has to make a number of important decisions.

[1] Following common practice, feminine forms will be used to refer to speakers, and masculine forms to refer to hearers.

First, she has to decide which information to include in the utterance, i.e., she needs to select the content of the message to be expressed. Once relevant concepts have been selected, these have to be put into a grammatical structure. Given that speech proceeds serially, this structure ultimately has to map on a linear order of words. That is, one thing has to be mentioned before another. Hence, a speaker needs to *choose a concept that will be referred to first*. Although languages may have grammatical restrictions on what types of entities an utterance can start with (e.g., the subject), there is a general tendency for entities that are conceptually highly salient (e.g., topical or animate/agentive) to be mentioned first (e.g., Van Bergen, 2011; Levelt, 1989; Tomlin, 1986).

Second, the speaker has to decide which linguistic form she is going to use to refer to a certain concept. That is, she has to *choose a referring expression*. Language provides an in principle infinite number of possible ways to refer to something, ranging, for example in English, from very elaborate expressions such as full definite descriptions with modifiers (e.g., *the large old-fashioned red stapler with the little scratch on the top*) to very short ones such as pronouns (e.g., *it*). In fact, given that the association between meaning and linguistic forms is largely arbitrary and based on convention (de Saussure, 1916/1959), any expression might do the job. However, there are regularities that make a certain type of expression more likely to be used in a certain situation. For example, speakers generally find it important that their expression can be interpreted correctly by the hearer. This will prevent them from saying, e.g., 'could you hand me the pineapple' or 'the sasamajah', when referring to the stapler, unless speaker and hearer have made an agreement on this way of referring to that particular object (e.g., Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986). In addition, referring expressions tend to become shorter when the same object is referred to multiple times (e.g., Clark & Wilkes-Gibbs, 1986). In theories of reference (e.g., Ariel, 1990; Chafe, 1994; Givón, 1983), speakers are commonly believed to choose referring expressions in such a way that these signal to the addressee how easily the referent can be accessed from memory, and hence aid the addressee in retrieving the correct antecedent. In general, the more accessible a referent is, the more reduced the expression referring to it will be.

Using language production experiments, conducted in Dutch, the research presented in this dissertation investigates factors that may influence the two types of referential choice mentioned above: the choice of a referent that will be mentioned first and the choice of a referring expression for that referent. Notably, it explores how

linguistic factors (grammatical function and lexical animacy) interact with non-linguistic factors (visual foregrounding and perceptual animacy) and speaker-internal factors (uncertainty and cognitive load). Regarding the choice of referent for first mention, it is investigated how and to what degree these factors influence whether an entity becomes the subject of the sentence, which is often the first-mentioned element in Dutch. The focus of this dissertation is however on the choice of referring expression, for which interactions between linguistic and non-linguistic or speaker-internal factors have not been studied much. Here, the area of interest is the choice of a particular type of referential form, rather than the selection of semantic content to include in a noun phrase (e.g., how speakers choose between 'the large stapler', 'the red stapler', and 'the large red stapler'). In particular, it is investigated how and to what degree the factors mentioned above influence speakers' choices for pronouns and full noun phrases in discourse.

Pronouns are defined as both phonologically and semantically attenuated expressions (e.g., Almor, 1999; Givón, 1976), i.e., they are typically short expressions that only carry some general semantic features, such as number, gender, and person. They can also be syntactically and/or prosodically restricted, such as reduced pronouns in Dutch, which cannot be stressed. This dissertation is only concerned with third person singular personal pronouns, both full and reduced, such as *hij*/*ie* 'he', *zij*/*ze* 'she', and *het* 'it', although Chapter 4 also discusses demonstrative pronouns such as *die* 'that' and *deze* 'this'. Expressions that contain a noun, possibly supplemented by determiners and modifiers, are referred to as full noun phrases. In the context of this dissertation, the term 'full noun phrase' usually means definite noun phrase, such as *de man* 'the man' or *de vrouw* 'the woman' (as opposed to indefinite noun phrase).

Before moving on to the main research questions of this dissertation, the next sections will provide a theoretical background on the notion of *accessibility*, which is generally assumed to drive referential choices in language production.

### 1.2. Accessibility and related terms

Both the choice of referent for first mention and the choice of referring expression have been related to the degree of activation of the conceptualizations, or mental representations, of referents in memory. The higher the activation of a certain

representation, the more likely it is to appear early in the linguistic structure, and the higher the likelihood that the expression referring to it is more attenuated (e.g., Levelt, 1989). This activation status has been described with a variety of terms, such as *accessibility* (Ariel, 1990; Bock & Warren, 1985), *salience* (Osgood, 1971; Sridhar, 1988), *cognitive status* (Gundel, Hedberg, & Zacharski, 1993), *givenness* (Chafe, 1976; Gundel et al., 1993; Prince, 1981), *topicality* (Givón, 1983) and *focus of attention* (Grosz, Joshi, & Weinstein, 1995), each with slightly different assumptions and viewpoints.

Some of these terms, such as givenness and topicality, emphasize the importance of information structure in the discourse. For example, when a referent in a discourse was the topic of the preceding sentence (with topic being defined as what the sentence is about; Reinhart (1982)), its representation in memory is likely to be highly activated. Other terms, such as focus of attention and cognitive status, emphasize the importance of cognitive capacities. For example, it seems likely that those referents that are attended to are more activated, since they may be actively maintained in memory (Foraker & McElree, 2007). To remain implicit as to the source of the activation, the more general term accessibility is used throughout this dissertation to refer to the ease of activation of mental representations in the memories of speakers and hearers, whatever the cause. For the sake of brevity, 'the accessibility of a referent' will be often used throughout this dissertation as shorthand for 'the accessibility of the mental representation of a referent'.

Crucially, this notion of accessibility concerns activation of non-linguistic representations, rather than activation of lexical items in the mental lexicon (cf. Arnold, 2010). To distinguish activation of non-linguistic representations from activation of lexical items, Bock and Warren (1985) speak of *conceptual accessibility*, which they define as "the ease with which the mental representation of some potential referent can be activated in or retrieved from memory" (p. 50), as opposed to *lexical accessibility*, which refers to "the ease with which the representations of word forms can be recovered from memory" (p. 52). In this dissertation, the term accessibility is used to refer to conceptual accessibility, unless explicitly specified otherwise. Furthermore, the term salience is reserved for properties of the referent itself rather than of its representation in memory. These properties can be linguistic, as when the referent is mentioned in a prominent or non-prominent syntactic position,[2] or non-linguistic (e.g.,

[2] Depending on the language, prominent syntactic positions include the subject, topic or preverbal position, for example, while object, focus and postverbal positions may be considered non-prominent.

perceptual), as in the size or color of the physical object that is referred to. They can also be determined by the context, such as the preceding discourse or the physical environment, or they can be intrinsic to the referent, such as animacy. Finally, the terms topicality and givenness are taken to denote factors that contribute to an entity's (linguistic) salience, while focus of attention is used as a speaker- or hearer-internal factor that might influence accessibility directly. Of course, these notions are all closely related, and in practice it might be difficult to keep them apart. For example, topical or given information is highly salient, by which it will attract attention, which in turn will increase the accessibility of the corresponding mental representations. However, on a theoretical level it is important to distinguish the *cause* of a low or a high accessibility of a mental representation from the degree of accessibility itself.

Thus, accessibility is thought to be a determining factor both in the choice of referent for next mention and the choice of referring expressions. However, research has revealed differences in how exactly accessibility affects these choices. Notably, the two types of referential choice may be affected by different factors (e.g., Fukumura & Van Gompel, 2010; Kehler et al., 2008; Stevenson, 2002; Stevenson, Crawley, & Kleinman, 1994), and they may differ in the degree to which accessibility refers to the referent's activation in the speaker's or the addressee's memory (e.g., Arnold, 2008). I return to this issue in Section 1.5. The next two sections discuss relevant literature on the role of accessibility in the choice of referent for first mention (Section 1.3) and in the choice of referring expression (Section 1.4).

### 1.3. Effects of accessibility on the choice of referent for first mention

When people speak (or write), seemingly unordered thoughts and concepts have to be put into the linear order that characterizes language. As a first decision, speakers have to choose a *starting point* from which their utterance is going to unfold (MacWhinney, 1977). The question is how they do this. It has been proposed that the decision of what information to present first and what later is affected by the accessibility of the concepts to be expressed (e.g., Bock & Warren, 1985). Indeed, it has frequently been observed that concepts with properties that make them highly salient are produced earlier in a linguistic utterance than less salient concepts. For example, referents that are perceptually salient, more animate, more imageable, or constitute given information, have typically been found to have a higher likelihood to be produced

early in the sentence (e.g., Bock, 1982; Bock & Irwin, 1980; Bock & Warren, 1985; Ferreira & Yoshita, 2003; Flores d'Arcais, 1975; Osgood & Bock, 1977; Prat-Sala & Branigan, 2000; Sridhar, 1988; Tomlin, 1997).

While the relation between accessibility and the positioning of concepts in the sentence may be direct, such that what is most accessible is produced first, it may also be mediated by grammatical function or topichood. For English, for example, it has been found that the most accessible concept is typically made the subject (e.g., Bock & Warren, 1985; McDonald, Bock, & Kelly, 1993). This suggests that accessibility determines which entity becomes the subject of the sentence, which in turn is preferably produced in the sentence-initial position, but that it does not determine sentence position directly. However, subject and sentence-initial position are highly confounded in English, which makes the exact relation between accessibility and sentence position unclear. In languages in which word order is more free, such as Greek (Branigan & Feleki, 1999), German (Kempen & Harbusch, 2004), Hungarian (É. Kiss, 2002), Italian and Spanish (Brunetti, 2009), accessibility has been found to affect sentence position independently of grammatical function. However, in such languages, accessibility may still affect the likelihood that something becomes a topic, and hence that it will occupy the topic position, which is often the first position in the sentence (Lambrecht, 1994). In a study of the Algonquian language Odawa, Christianson and Ferreira (2005) were able to disentangle effects on both grammatical function and topichood from those on linear order by looking at different verb forms in that language. They found that accessible entities in Odawa were not directly promoted to the sentence-initial position, but were given prominent syntactic functions via the priming of a particular syntactic structure.

In this dissertation, the question whether the influence of accessibility on the choice of referent for first mention is direct or indirect, via grammatical function and/or topicality, is not dealt with. Although in Dutch, the language under investigation, both starting a sentence with the subject and starting a sentence with the topic are important preferences (e.g., Bouma, 2008; Vogels & Van Bergen, 2013), first mentioned entities in the studies presented in this dissertation are mostly also subjects. Therefore, only the effect of accessibility on the likelihood that a referent will be the subject is investigated.

If we accept that accessibility affects the choice of referent for first mention, whether directly or indirectly, the next question is which factors determine accessibility, how these factors are related and how they interact with each other (cf. Arnold, 2010).

Givón (1976) proposes that different saliency factors, such as animacy, agency and givenness, combine to form a hierarchy of topicality. Since people tend to talk about animate agents, for example, such entities are likely to be the topic of the sentence, and hence to occur in a prominent (e.g., sentence-initial) position. This also relates to the *predictability* of a referential act: What people tend to talk about is expected to be mentioned next and therefore accessible for the hearer (Arnold, 2001; Givón, 1983). On the other hand, predictable entities may also be postponed to a less prominent position, due to a preference to start an utterance with the most important (i.e., most newsworthy) information (Givón, 1983; 1988; Gundel, 1988).

Alternatively, what these saliency factors may have in common is that they attract attention (e.g., Gleitman, January, Nappa, & Trueswell, 2007; Myachykov, Garrod, & Scheepers, 2009; Tomlin, 1997). Perceptual attention may be captured by, e.g., large, foregrounded, animate or moving objects (e.g., Flores d'Arcais, 1975; Mazza, Turatto, & Umiltà, 2005; New, Cosmides, & Tooby, 2007; Pratt, Radulescu, Guo, & Abrams, 2010). In a discourse, elements in a prominent syntactic function (e.g., subject) may be in the focus of attention (e.g., Grosz et al., 1995). Because what is attended to is easier to retrieve, it is more likely to be talked about first.

Different sources of accessibility may also interact. Prat-Sala and Branigan (2000) distinguish two types of accessibility: A referent's *inherent accessibility* refers to activation in memory caused by its intrinsic properties, such as its animacy or concreteness, which are assumed to be stable across contexts. Within a discourse, this inherent activation can be supplemented by the referent's *derived accessibility*, a temporary level of activation caused by the salience of the referent in the discourse, such as whether it is given or topical. Thus, a referent's derived accessibility adds to its inherent accessibility. If the two types of accessibility run counter to each other, such as when the referent is inanimate but given, derived accessibility may override inherent accessibility if the context is strong enough (Prat-Sala & Branigan, 2000). Van Nice and Dietrich (2003b) also found an interaction between inherent (animacy) and derived (thematic role) accessibility, but only when speakers had to speak from memory, as opposed to describing pictures in view. They argued that in that case speakers process information from multiple referents simultaneously, allowing different types of information to interact.

An important question is whether a speaker chooses a referent for first mention because it is highly accessible for herself, or because she assumes it is highly accessible (or wants it to become highly accessible) for her addressee. Speakers may put a certain

element in the sentence-initial position to invite the addressee to pay attention to that element and use it to store subsequent information (e.g., the utterance 'Vladimir tickled Barack' should be stored under 'things that Vladimir did', while 'Barack was tickled by Vladimir' is probably stored under 'things that happened to Barack'; Givón, 1988; Levelt, 1989). Alternatively, speakers might produce those word orders that are easiest to interpret for the hearer (Hawkins, 1994).

Despite this possibility, conceptual accessibility is generally taken to be speaker-oriented, i.e., it is assumed to involve the activation of mental representations in the speaker's rather than the addressee's memory (e.g., Bock & Warren, 1985; Prat-Sala & Branigan, 2000). If we assume that language production proceeds incrementally (e.g., Kempen & Hoenkamp, 1987; Levelt, 1989), speakers start producing an utterance before the planning of that utterance is completed. Because highly accessible referents are more easily retrieved from memory, they are subsequently mentioned earlier in the sentence. Indeed, studies have found that visual attention of the speaker influences order of mention (e.g., Gleitman et al., 2007; Tomlin, 1997). Gleitman and colleagues, for example, presented participants with simple scenes (e.g., of a dog chasing a man), and found that these scenes were described with active ('the dog chases the man') or passive ('the man is chased by the dog') sentences, depending on the location of a not consciously noticeable attentional cue (a black square, presented very briefly either on the dog or on the man). In addition, speakers do not seem to avoid ambiguities for their addressees when producing certain syntactic structures (Arnold, Wasow, Asudeh, & Alrenga, 2004). These findings suggest that the choice of referent for first mention is influenced by speaker-internal constraints, rather than by addressee-oriented processes.

Central to this dissertation is the question whether the non-linguistic and speaker-internal factors that have been found to affect the choice of referent for first mention, such as animacy, visual salience and speaker attention, also affect the choice of referring expression. This is the topic of the next section.

### 1.4. Effects of accessibility on the choice of referring expression

Several theoretical accounts have been proposed to answer the question how speakers choose between different types of expressions to refer to something, ranging from definite and indefinite noun phrases to pronouns and zero anaphora. In contrast to

research on the choice of referent for first mention, theories on the choice of referring expression have mainly concentrated on discourse factors such as givenness and topicality. Below, the most important accounts, which are similar in a number of respects but differ in some of their assumptions, are briefly discussed.

*1.4.1. Ariel (1990)*

In Ariel's theory of accessibility (Ariel, 1990; 2001), speakers choose referring expressions such that these provide the addressee with information about the current activation state of the referent in the discourse. In that way, addressees know where in memory they have to look for the mental representation to be retrieved. The general rule is that the more accessible a referent is deemed, the shorter and more attenuated (either phonologically or semantically) the referring expression will be. Conversely, the longer and more informative the referring expression is, the lower the degree of accessibility it codes will be. Ariel (1990) distinguishes three main types of referring expressions according to the degree of accessibility that they code. Firstly, expressions such as definite descriptions and proper names are low accessibility markers: They indicate that the memory representation of the referent is probably not activated. Secondly, demonstrative noun phrases and demonstrative pronouns code an intermediate degree of accessibility and hence are medium accessibility markers. Finally, highly reduced expressions such as pronouns, clitics (i.e., elements that are phonologically bound to another word) and zero anaphora (i.e., empty referring expressions, as in 'Mandy was tired and Ø fell asleep') make up the high accessibility markers. These expressions are used when the speaker has reason to believe that the hearer currently has a highly activated representation of the referent.

Thus, accessibility in this view refers to a property of a referent in a discourse, which a speaker marks for the addressee by using a certain linguistic form. According to Ariel (1990), accessibility is influenced by different discourse factors, such as topicality, grammatical function, recency, frequency, competition and predictability. For example, a referent that has recently been mentioned is likely to be referred to with a high accessibility marker. Hence, in the second sentence of the Dutch example in (1a) a pronoun will generally be preferred to refer to Fons, who is mentioned in the directly preceding sentence.[3] Here, repeating the name would give rise to the implication that the discourse contains two people named Fons. However, in (1b) a name would be preferred over a pronoun to refer to Fons, despite the fact that Fons is

[3] Indexes mark coreferentiality. Hashes (#) mark low acceptability given the context.

still the most recently mentioned entity. This is because there is a competing entity, Emiel, which is mentioned in subject position and which is more topical (i.e., the sentence is more about Emiel than about Fons). These factors also contribute to accessibility.

(1) a. *Fons$_i$ was in de tuin aan het werken. Plotseling werd*
F. was in thegarden on thework suddenly became
*{hij$_i$/#Fons$_i$} geraakt door een zwiepende tak.*
he/F. hit by a swishing branch
'Fons$_i$ was working in the garden. Suddenly, {he$_i$/#Fons$_i$} was hit by a swishing tree branch.'

b. *Emiel was Fons$_i$ aan het helpen in de tuin. Plotseling werd*
E. was F. on thehelp in thegarden suddenly became
*{#hij$_i$/Fons$_i$} geraakt door een zwiepende tak.*
he/F. hit by a swishing branch
'Emiel was helping Fons$_i$ in the garden. Suddenly, {#he$_i$/Fons$_i$} was hit by a swishing tree branch.'

To show that it is not always the case that pronouns refer to the highest grammatical function (i.e., the subject) in the preceding sentence, consider the example in (2). Here, *hij* 'he' most likely refers to Constantijn, despite Hans being the subject of the preceding sentence, because it is likely that the second sentence is providing the reason why Constantijn was admired.

(2) *Hans was trots op Constantijn$_i$. Hij$_i$ kon in 20 seconden een hele*
H. was proud on C. he could in 20 seconds a whole
*taart verorberen.*
cake devour
'Hans was proud of Constantijn$_i$. He$_i$ could devour an entire cake in 20 seconds.'

This suggests that predictability is another important factor in determining accessibility (but see Section 1.5.2 below for findings that contest this). In sum, according to Ariel, none of these factors alone (e.g., grammatical function, competition,

predictability) can explain the variation in the use of referring expressions, but the complex notion of accessibility can.[4]

*1.4.2. Givón (1983)*

Other theoretical accounts explicitly focus on a single discourse factor as the main determinant of the degree of activation of referents in memory, but stretch it in such a way that it can cover the range of variation in referring expressions. Givón (1983) relates the use of different types of expressions to different degrees of *topic continuity*. Topic continuity refers to whether the same topic (i.e., what the sentence is about) is maintained in the preceding discourse, and whether it will *persist* in the subsequent discourse. Hence, it is a combination of the recency and the predictability of topical elements. Highly continuous topics are both recently mentioned and likely to be mentioned again. Therefore, they are more likely to be referred to with attenuated expressions such as pronouns. Topics with low continuity are either new in the discourse or not persistent, and will therefore be more likely to be referred to with elaborate expressions such as full noun phrases.

While acknowledging that many more factors may play a role, Givón argues that the concrete, measurable discourse factors underlying topicality (i.e., recency and predictability) can explain a significant part of the variation in referential forms. As in Ariel's theory, topicality forms a continuum, with a certain expression coding a certain part of the scale. However, cross-linguistically, this coding is only fixed in relation to other expressions. That is, a certain type of expression (say, a pronoun) may code some part of the topic continuity scale in one particular language, but this need not be the same part in another language. Yet, in no language does a pronoun code a lower position on the scale than the types of expression below it (say, demonstratives and full noun phrases).

*1.4.3. Gundel, Hedberg, and Zacharski (1993)*

In contrast to the continuous scales of Ariel (1990) and Givón (1983), Gundel et al. (1993) propose a discrete hierarchy of six cognitive statuses, which relate to the givenness of mental representations in the addressee's memory. Although the term givenness suggests that a referent's cognitive status is determined by whether the entity is given or new information in the discourse (e.g., Prince, 1981), it is intended as

[4] Still, Ariel (2001) notes that there may be some exceptions that have to be explained by other factors, such as contrastiveness.

a psychological notion, referring to what the addressee is currently focusing on (as believed by the speaker), whether related to the preceding discourse or not. By using a certain referring expression, a speaker signals to the addressee where or how he should mentally access the referent. For example, when a speaker assumes that the addressee already has a representation of a certain referent in memory, this licenses the use of a definite expression. If this representation is not only assumed to be present but also to be in the focus of attention, the use of a pronoun is appropriate. The cognitive statuses are said to be implicationally related, such that the use of a referential form to signal a certain status implies that all lower statuses have been met as well. Therefore, less attenuated expressions can in principle also be used to refer to entities in the focus of attention. Pragmatic constraints will however encourage speakers to be maximally informative, and discourage them to use expressions that are more elaborate than necessary (e.g., Grice, 1975).

*1.4.4. Chafe (1994)*

Chafe (1994) also relates the choice of referring expression to cognitive statuses. He limits the number of statuses to three: active, semiactive, and inactive (although the boundaries between those may be fuzzy). Active information is information that is in the focus of attention, while inactive information is unattended or unconscious. Semiactive information is somewhere in between, in the periphery of attention. What elements are active in the addressee's mind is not only determined by what information the speaker has brought forward, but also by the physical context, world knowledge, inferences, and shared knowledge between speaker and addressee (e.g., Chafe, 1994; 1996; Clark & Bangerter, 2004; Clark & Haviland, 1977; Gundel et al., 1993; Prince, 1981).

*1.4.5. Centering theory*

A formalization of the relation between focus of attention and referential form is provided by *centering theory* (Grosz, Joshi, & Weinstein, 1983; 1995). According to centering theory, each utterance in a discourse segment (Grosz & Sidner, 1986), except for the first one, has a single entity that is currently in the focus of attention, which is associated with how central that entity is in the local discourse. This entity is known as the backward-looking center ($C_b$). The $C_b$ relates to one of several forward-looking centers ($C_f$) in the previous utterance. The forward-looking centers are a partially ordered list of entities in an utterance, which are ranked by their discourse salience.

This salience is primarily determined by the entity's surface position and syntactic function, such that subjects rank higher than direct objects, which in turn rank higher than oblique objects (Gordon, Grosz, & Gilliom, 1993; Grosz et al., 1995). The highest ranked $C_f$ that also occurs in the next utterance is the $C_b$ of that utterance. In other words, what is in the focus of attention in a given utterance is determined by whether it was mentioned in a prominent position (e.g., sentence-initial or subject position) in the preceding utterance.

In interpreting a discourse, addressees have to make inferences about the relations between consecutive utterances. One of the assumptions in centering theory is that speakers seek to produce a maximally coherent discourse to minimize these inferences. To this end, they try to avoid too many shifts to a different backward-looking center across utterances. Speakers are also assumed to choose certain referring expressions to signal whether they continue to talk about the same thing: If any entity in the current utterance is pronominalized, this should at least be the backward-looking center. This means that, according to centering theory, pronouns are used to refer to the most discourse salient entity, but nothing prevents other entities from being pronominalized as well. In addition, the account also allows for a situation in which no previously mentioned entity is pronominalized at all. The assumptions of centering theory have been partly confirmed by both psycholinguistic experiments and corpus research (e.g., Brennan, 1995; Gordon et al., 1993; Poesio, Stevenson, Di Eugenio, & Hitzeman, 2004).

*1.4.6. Computational models of referring expression generation*

Formalisms of referring expression production such as centering theory have also been used as a basis for computational models of resolving and generating referring expressions in discourse (e.g., Brennan, Friedman, & Pollard, 1987). In addition, algorithms for generating pronouns in texts have been developed, which also take into account factors such as the antecedent's ontological type (Strube & Wolters, 2000) and global discourse structure (Callaway & Lester, 2002; McCoy & Strube, 1999). However, most models of referring expression generation (REG) focus on content selection for initial definite descriptions used for object identification outside of a discourse context. For example, the Incremental Algorithm (Dale & Reiter, 1995) is able to determine which and how much information should be included in an expression to result in correct identification of the referent, as in the choice between *the man in the suit*, *the man on the left*, or *the man in the suit on the left* (Krahmer & Van Deemter, 2012).

Krahmer and Theune (2002) propose an extension of the Incremental Algorithm such that it can also handle references in discourse. Instead of generating an expression that minimally distinguishes the target referent from its distractors, their algorithm chooses an expression based on the salience of the possible referents. As in centering, salience is based on the syntactic prominence of the entities in the context. Each entity receives a weight value between 0 and 10, which decreases with every utterance in which it is not mentioned. In this way, the algorithm can produce underspecified expressions for salient entities. For example, the single most salient referent in the set of possible referents is referred to with a pronoun, which was found to be in accordance with the preferences of human participants. Recently, the GREC challenges program (Generating Referring Expressions in Context; Belz, Kow, Viethen, & Gatt, 2010) has started to evaluate systems that generate referring expressions in discourse, including pronominal expressions. One of the aims of these systems is to produce human-like references within a context, making use of psycholinguistic data.

### 1.5. The prevalence of the role of the linguistic context and of the addressee

The frameworks discussed above (including the computational models) all share the idea that in a discourse, some entities are focused on more than others (both by the speaker and the addressee), and that this has an impact on the choice of referring expression. In each case, the degree of accessibility (or topicality/givenness/focus of attention) is presented as a property of mental representations, which is influenced by, but by no means identical to, the salience of referents in the preceding discourse. However, although it is acknowledged that referents that have not been mentioned previously can still be accessible, for instance from the physical context (e.g., '*that woman* over there') or from world knowledge (e.g., '*the king* will visit my hometown tomorrow'), the focus in research on the choice of referring expressions has been on the influence of the discourse context. This has been considered the most important factor driving the activation of mental representations. For example, Ariel (2001) claims that:

> [...] it is the discoursal rather than the physical salience of the entities involved which determines the degree of accessibility assigned to particular mental representations [...]. Although the physical context does affect the discourse model

> of the speakers, mental representations are a direct product of our discourse model only. (Ariel, 2001, p. 31)

In research on accessibility, referring expressions have thus been investigated mainly as anaphors, i.e., expressions that have an antecedent in the preceding discourse (or in the upcoming discourse in the case of cataphors). Hence, factors affecting accessibility have been primarily sought in properties of the antecedent. Discourse factors that have been identified in both psycholinguistic experiments and corpus studies as influencing the accessibility of the antecedent include, among others, recency (e.g., Clark & Sengul, 1979), topicality (e.g., Givón, 1983), first mention (e.g., Gernsbacher & Hargreaves, 1988), grammatical function (e.g., Brennan, 1995; Gordon et al., 1993), syntactic parallelism (e.g., Arnold, 1998), competition (e.g., Ariel, 2001; Arnold & Griffin, 2007), protagonisthood (e.g., Karmiloff-Smith, 1981; Morrow, 1985), episode shifts (e.g., Anderson, Garrod, & Sanford, 1983; Vonk, Hustinx, & Simons, 1992), and thematic role (e.g., Arnold, 2001; Stevenson et al., 1994). Although some of these factors (especially the last three) may also apply to the non-linguistic context, they have primarily been investigated in linguistic contexts.

Another common assumption in theories of reference production has been that speakers choose referring expressions for their addressees. According to both Ariel (1990) and Gundel et al. (1993), for example, speakers choose referring expressions based on their assumptions about the activation status of the referent in the addressee's memory. In communication, speakers are supposed to adhere to the principle of optimal design (Clark, Schreuder, & Buttrick, 1983): They design their utterances to be in accordance with what they believe their addressee knows (what is in common ground between speaker and addressee), such that the addressee can easily pick out the correct referent. Speakers are also expected to abide by the Gricean Maxim of Quantity, such that they choose referring expressions that are as informative as required, but not more informative than required (Grice, 1975). For example, when the context contains two possible referents of the same gender, using a pronoun (e.g., *he* or *she*) to point out one of these referents would, at least in English, result in ambiguity for the addressee. Therefore, speakers are predicted to avoid this type of ambiguity by choosing more specific referring expressions, something that has also been found experimentally (e.g., Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Arnold & Griffin, 2007; Fukumura, Van Gompel, & Pickering, 2010).

These two assumptions, i.e., referring expressions are chosen based on a model of the discourse and they are tailored for an addressee, are both reflected in the account of Brennan and Clark (1996). They argue that while factors such as perceptual salience may influence the choice of a referring expression, what is most important is whether the referent has been mentioned recently or frequently in the discourse. In addition, referring expressions are established in interaction with addressees. Thus, in the classic view on how speakers choose a particular referential form, accessibility refers to the degree of activation of mental representations in the addressee's memory, as assumed by the speaker. This assumed activation is mainly determined by whether the representations are believed to be in common ground between the speaker and the addressee, to which the discourse context (i.e., whether and how the referent has been mentioned before) makes the greatest contribution.

Non-linguistic factors, such as perceptual salience and intrinsic properties of referents, have typically not been taken into account in traditional theories of reference production. Still, perceptually and conceptually salient entities are likely to attract attention (e.g., Coco & Keller, 2010; Henderson & Ferreira, 2004; New et al., 2007; Pratt et al., 2010), and may therefore influence referent accessibility (Arnold & Griffin, 2007). Physical presence is an important source of the referent's accessibility (e.g., Clark & Marshall, 1981). For example, expressions such as unheralded pronouns (pronouns without a linguistic antecedent) and deictics (e.g., *that one*, often accompanied by a pointing gesture) are dependent on the configuration of objects in the physical environment of the interlocutors (e.g., Clark et al., 1983; Greene, Gerrig, McKoon, & Ratcliff, 1994; Jarvella & Klein, 1982; Piwek, Beun, & Cremers, 2008). Indeed, it has been found that the physical context affects the production of referring expressions (e.g., Beun & Cremers, 1998; Ferreira, Slevc, & Rogers, 2005; Fukumura et al., 2010; Osgood, 1971; Sedivy, 2003; Sridhar, 1988). In addition, there is evidence that higher-level conceptual properties of referents, such as animacy, individuation and concreteness, affect the choice of referring expressions (e.g., Brown-Schmidt, Byron, & Tanenhaus, 2005; Dahl & Fraurud, 1996; Fukumura & Van Gompel, 2011; Maes, 1997; Maes & Noordman, 1995; Yamamoto, 1999). Yet, little is known about how non-linguistic factors interact with linguistic factors in referential choices.

It has also become clear that speakers are not always optimally designing referential forms for their addressees (e.g., Gann & Barr, 2012; Keysar, Barr, & Horton, 1998). For example, speakers often overspecify their referring expressions (e.g., Arts, 2004; Engelhardt, Bailey, & Ferreira, 2006; Koolen, Goudbeek, & Krahmer, 2013). In a

referential description task described in Koolen et al. (2013), speakers often used color information when providing definite descriptions for objects in a scene, sometimes even when all objects in the scene had the same color. This information would be completely uninformative for an addressee trying to identify the object that is being referred to. One probable explanation for the fact that speakers do not always take the addressee's communicative needs into account is that speaker-internal factors, such as restrictions on memory and attention resources and principles of economy, also play an important role in the choice of referring expressions (e.g., Arnold, 2008; Arnold & Griffin, 2007; Fukumura & Van Gompel, 2012; Hendriks, Koster, & Hoeks, 2013; Horton & Keysar, 1996). The question of how much in reference is addressee-oriented and how much is attributable to speaker-internal factors is currently subject of debate in the literature (see Arnold, 2008 for a review).

To sum up the picture sketched so far, when speakers refer to an entity they appeal to a mental representation of the referent. This representation is assumed to have a certain degree of activation, which can be determined by many different factors, either linguistic (e.g., structural properties of a previous mention of the referent in the discourse) or non-linguistic (e.g., the referent's visual salience); and either contextual (determined by the preceding discourse or physical environment) or intrinsic (e.g., animacy). The higher the activation, the more accessible the referent, i.e., the easier it can be retrieved from memory. Assuming that spoken language production proceeds incrementally, the element that is retrieved quickest is also produced first (within grammatical constraints). Thus, more accessible referents are more likely to be chosen for first mention in an utterance. Because highly accessible referents do not need a lot of linguistic encoding to be retrieved by an addressee, speakers are more likely to use more attenuated expressions, such as pronouns, to refer to such referents. Based on the research reviewed here, the question arises what exactly underlies the accessibility effects on both the choice of referent for first mention and the choice of referring expression. This question is discussed in the next section.

### 1.6. What underlies effects of accessibility?

So far, accessibility has been defined as the degree of activation of a referent's mental representation, which may be influenced by many different factors. It is not yet clear what it is that underlies this variety of factors, binding them together. For an answer

to this question, it is interesting to compare reported accessibility effects on the choice of referent for first mention and on the choice of referring expression.

*1.6.1. Comparing accessibility effects on different referential choices*

In Sections 1.3 and 1.4, it was discussed how accessibility is assumed to affect both the choice of referent for first mention and the choice of referring expression. From this discussion, an interesting incongruence emerges between the two referential choices. Research on the choice of referent suggests that salient entities are preferred to be mentioned first in the sentence because they are in the speaker's focus of attention. Hence, accessibility is assumed to be influenced by factors that capture the attention of the speaker (e.g., visual cues, animacy; Gleitman et al., 2007; Prat-Sala & Branigan, 2000). Accessibility effects on the choice of referring expression, however, are mostly attributed to the need to be clear for the addressee (Ariel, 1990; Gundel et al., 1993). In this view, speakers attenuate their expressions when they believe that the referent is in the focus of attention of their addressee. Here, accessibility refers to the assumed activation of referents in the mind of the addressee. Indeed, research on the choice of referring expression has focused on the influence of information that is generally available for both speaker and addressee, such as the preceding discourse (e.g., givenness, topicality, grammatical function; Brennan, 1995; Chafe, 1976; Clark & Marshall, 1981; Gundel et al., 1993; Prince, 1981).

Thus, how speakers choose a referent for first mention has often been explained in terms of the speaker's focus of attention, or the topic the speaker wants to talk about; conversely, the choice of referring expression has been attributed to the speaker's assumption about the addressee's focus of attention, or what the addressee expects the speaker to be talking about. However, if both referential choices (what to refer to first and how to refer to it) are determined by the same concept of accessibility, one would predict that what makes a referent accessible will increase the probability that it will be mentioned first as well as the probability that it will be referred to with an attenuated expression. This implies that what is accessible is what is most likely to be talked about next, i.e., most predictable.

*1.6.2. Accessibility as predictability*

As mentioned above, Givón (1983) argues that more accessible topics are more likely to be talked about again and hence are more predictable. Because predictable entities need less explicit encoding, they can be referred to with more attenuated expressions

(see also Jaeger, 2010). A similar view has been proposed by Arnold (1998; 2001) in her *Expectancy hypothesis*: Accessibility correlates with the probability that a referent will be mentioned again (Arnold, 2008). Hearers make predictions about what the speaker is going to say based on the preceding discourse and on their experience with what types of entities (e.g., subjects, animates) speakers are likely to keep on talking about. Hearers will be more likely to perceive these entities as being accessible for the speaker, who will be more likely to use attenuated expressions to refer to them (for speaker-internal reasons). In this way, the speaker's choice of referring expression aids the addressee's comprehension (whether so intended or not), since it signals whether the information was predictable or not. Evidence for this view was found by Arnold (2001), who used a story continuation task to elicit references to entities with either a source or a goal role, counterbalanced for grammatical function (e.g., *Lisa*$_{\text{SOURCE}}$ *gave the leftover pie to Brendan*$_{\text{GOAL}}$ or *Marguerite*$_{\text{GOAL}}$ *caught a cold from Eduardo*$_{\text{SOURCE}}$ *two days before Christmas*). In their continuations, participants were more likely to refer to the goal entity than to the source entity. While grammatical subjects were most likely to be pronominalized, participants were also more likely to use pronouns to refer to the goal entity than to the source entity, at least for non-subject referents. Arnold (2001) concluded that goal entities are more accessible than source entities because they are more likely to be mentioned next. Hence, speakers will (inadvertently) signal this accessibility to addressees by using attenuated expressions.

Other researchers question the view that what the speaker refers to first and how she refers to it are both determined by the predictability of a referential event. For example, like Arnold (2001), Stevenson et al. (1994) found in a story completion study that participants were most likely to continue the story with the character that was in a goal role (e.g., *John*$_{\text{SOURCE}}$ *passed the comic to Bill*$_{\text{GOAL}}$), a patient role (e.g., *Joseph*$_{\text{AGENT}}$ *hit Patrick*$_{\text{PATIENT}}$), or either a stimulus or experiencer role (e.g., *Ken*$_{\text{STIMULUS}}$ *impressed Geoff*$_{\text{EXPERIENCER}}$) depending on the connective that followed the introductory sentence (e.g., *because* or *so*). However, the choice of whether to use a pronoun or not was only affected by the referent's syntactic position, with participants being more likely to use a pronoun when referring to the first mentioned character than to the second mentioned character.

More recently, additional evidence that the choice of referring expression is independent from the likelihood that the referent will be mentioned next was obtained by Rohde (2008; see also Kehler et al., 2008), Fukumura and Van Gompel (2010), Chiriacescu (2011) and Kaiser, Li and Holsinger (2011). These studies suggest

that the choice of referring expression may rather be driven by local discourse factors such as syntactic function and topichood. Kehler et al. (2008) offer a Bayesian approach to pronoun production and interpretation that captures the dissociation between the choice to pronominalize and the likelihood of next mention, presented in (3).

(3) $$P(referent|pronoun) = \frac{P(pronoun|referent)\,P(referent)}{P(pronoun)}$$

In this account, *P*(*referent*|*pronoun*) represents the probability that a pronoun refers to a certain referent. This probability is dependent on both the probability that given a certain referent, a pronoun is used to refer to it (*P*(*pronoun*|*referent*)) and the probability that that referent is mentioned (*P*(*referent*)). Crucially, *P*(*pronoun*|*referent*) is taken to be independent of *P*(*referent*), enabling both referential choices to be determined by different factors.

*1.6.3. Accessibility as a multiple-constraints factor*

The opposition between what speakers refer to first and how they refer to it raises the question what it actually means to say that a referent is accessible. Some researchers suggest that accessibility is only related to previous mention in the discourse, and not affected by predictability based on semantic factors (e.g., Fukumura & Van Gompel, 2010). Other researchers propose that accessibility is a multi-faceted notion, capturing different constraints that may interact and sometimes work against each other (e.g., Arnold, 1998). For example, Kaiser and Trueswell (2008) propose that referents can be accessible on different levels of representation. On the one hand, a referent can be accessible in a syntactico-semantic representation of the preceding sentence, which includes information about linguistic factors such as grammatical function and thematic role. On the other hand, it can be accessible in a mental model of the discourse, which involves a global representation about the event being described (Johnson-Laird, 1983; Kintsch & Van Dijk, 1978).

Alternatively, the opposition might be between the inherent accessibility of a referent and the additional accessibility it may receive from the discourse context (Prat-Sala and Branigan, 2000), or between a referent's accessibility for the speaker and its assumed accessibility for the addressee (e.g., Arnold, 2008; Bard et al., 2000; Galati & Brennan, 2010). Crucially, different referential choices may be either more or less sensitive to different accessibility-related factors (e.g., Brown-Schmidt et al., 2005;

Kaiser & Trueswell, 2008). For example, while the choice between pronouns and more elaborate referring expressions may be affected by discourse salience, the choice between full and reduced pronouns in languages such as Dutch and Estonian may be more related to contrastiveness (Kaiser, 2010; 2011; Kaiser & Trueswell, 2004). In the same vein, some factors may be more important for the choice of referent for first mention, while others are more important for the choice of referring expression. The present dissertation explores the degree to which different referential choices are affected by different factors.

### 1.7. Research questions

The aim of this dissertation is to gain a better understanding of the role of accessibility in referential choices. In particular, the research presented here investigates whether non-linguistic and speaker-internal factors, which have been argued to influence the choice of referent for first mention, also influence the choice for a particular referring expression, and if so, how they interact with linguistic factors. To what degree are different referential choices sensitive to different accessibility factors? The next four chapters present studies investigating interactions between different factors that are assumed to affect accessibility. They will deal with the following research questions:

- Do non-linguistic saliency factors, in particular visual foregrounding and perceptual animacy, affect the choice between pronouns and full noun phrases? (Chapters 2 and 3)
- How do these non-linguistic factors interact with linguistic factors (grammatical function and lexical animacy)? (Chapters 2 and 3)
- Do these linguistic and non-linguistic factors affect the choice of referent and the choice of referring expression differently? (Chapters 2 and 3)
- If the effect of animacy on pronoun use is due to the accessibility of the referent, does animacy also affect other referential choices, such as that between full and reduced pronouns? (Chapter 4)
- If the effect of animacy on pronoun use is due to the accessibility of the referent, does animacy have similar effects in language varieties that differ in the degree to which speakers tend to avoid gendered pronouns for inanimate nouns? (Chapter 4)

- Given that it is not yet clear whether a referent is highly accessible when it is in the speaker's focus of attention or when it is in the addressee's focus of attention, as estimated by the speaker, how does an increased cognitive load for the speaker influence the choice of referring expression? (Chapter 5)

### 1.8. Methodology

The studies in this dissertation employ several language production experiments, conducted in Dutch. The experiments presented in Chapters 2 and 5 are story completion experiments, which are commonly used in research on both reference production and reference comprehension (see, for example, Arnold, 2001; Fukumura et al., 2010; Kaiser & Trueswell, 2008; Stevenson et al., 1994). These tasks require participants to both produce and comprehend utterances, as they have to read or listen to the beginning of the story and then provide a continuation themselves. Chapter 4 also uses story completion, but in a more restricted way. Here, only a referring expression has to be filled in. Finally, Chapter 3 employs a retelling task, in which participants watch animated visual scenes and recount them afterwards. While the experiments in Chapters 2 and 4 involve no addressee, which is the case in most story completion experiments, participants in Chapters 3 and 5 tell their stories to another person. Whereas many previous story completion studies have used written tasks, the studies presented here are concerned with spoken language production. This is relevant because it is not necessarily the case that spoken reference production involves the same processes as written reference production.

The data obtained by the experiments are analyzed with mixed-effect models (e.g., Baayen, Davidson, & Bates, 2008), using the lme4 package in the R software program (Bates, Maechler, & Bolker, 2013). The use of mixed models allows the analysis of both between-participant and between-item variation in a single statistical model. This is an advantage to repeated measurements ANOVAs, widely used in psycholinguistics, which require the so-called F1 and F2 analyses to model this variation. Performing separate analyses for generalization over participants and items may lead to increased Type I error rates (Barr, Levy, Scheepers, & Tily, 2013; Clark, 1973). Since the studies presented here are concerned with categorical observations (e.g., whether a pronoun was used or not), we make use of *logit* mixed models, which can handle binomially distributed data, unlike ANOVA (Jaeger, 2008).

### 1.9. Overview

This section presents an overview of the remaining chapters of this dissertation. Chapters 2 to 5 are self-contained texts, which are based on articles either published or submitted for publication in peer-reviewed journals. Hence, some overlap between the individual chapters, and between those chapters and this introduction will occur. The author of this dissertation was the main researcher in all studies presented here.

The study presented in **Chapter 2** investigates the interaction between the salience of the referent in the perceptual context (visual foregrounding) and its salience in the linguistic context (grammatical function) in referential choices. Two spoken story completion experiments show that while visually salient referents are more likely to be referred to first, they are not more likely to be pronominalized. By contrast, linguistically salient referents (subjects) are more likely to be referred to with pronouns in a subsequent utterance, but they are not more likely to be referred to first. However, referents that have been introduced as the discourse topic *are* more likely candidates for first mention. These findings are taken as evidence that the choice of what to refer to first is mainly driven by the speaker's global representation of what is most important to talk about, which may be influenced by high-level factors such as protagonisthood as well as by low-level visual salience. Conversely, the local accessibility of individual referents is not affected by this, being primarily driven by structural properties of the linguistic context.

**Chapter 3** is concerned with the inherent salience of referents (animacy). Two animation-retelling experiments are presented in which again the interaction between a perceptual factor (perceptual animacy) and a linguistic factor (lexical animacy) was investigated. The results show that the two factors affect referential choices differently. Whereas entities that are perceived as animate are more likely to be pronominalized, they are not more likely to be mentioned first. Conversely, entities with animate names (but not necessarily perceived as animate) are more likely to be mentioned first, but not more likely to be pronominalized. These findings show that the perceptual context can influence the choice of referring expression, and that this effect should be dissociated from animacy effects on likelihood of next mention. The results are discussed in light of the time course of linguistic processing and task-specific dependencies.

**Chapter 4** zooms in further on effects of animacy on the choice of referring expression in Dutch. The aim of this chapter is twofold: Firstly, it examines the

predictions of theories of accessibility on a more fine-grained scale of referential forms, by looking at the distribution of full and reduced pronouns. Secondly, it investigates whether (lexically) animate entities are more likely to be pronominalized in Dutch, and to what degree this can be explained by a gender avoidance strategy rather than by accessibility. By comparing reference production of Dutch speakers from the Netherlands that lack grammatical intuitions about masculine and feminine nominal gender to Dutch speakers from Belgium that largely retain these intuitions, we are able to show that gender avoidance alone cannot explain the animacy effect. Hence, this effect is likely to be mainly driven by accessibility. However, an accessibility account cannot explain the distribution of full and reduced pronouns, which may instead be affected by the importance of information for the speaker or the addressee. This supports the view that choices for different types of referring expressions are driven by different factors.

The aim of **Chapter 5** is to investigate whether the choice of referring expression is influenced by a speaker-internal constraint, cognitive load. In two experiments, speakers completed stories for an addressee, which required them to refer to both salient and non-salient characters in a picture. In one condition, they also performed a memory task at the same time. The effect of this increased cognitive load was that pronouns became used more often, at least for non-salient referents. This was the case both when the referent's salience differed between the speaker's and the addressee's perspectives and when all discourse information was shared. This result suggests that accessibility is not identical to speaker attention. At the same time, speakers did not seem to take into account the referent's accessibility for the addressee. We propose that speakers use their own discourse model as a proxy for their addressee's, but that they have more difficulties taking this discourse model into account when they experience an increased cognitive load, resulting in more egocentric (economical) expressions.

Finally, **Chapter 6** provides a general discussion of the main findings, relating them to the theories of reference production discussed above, and specifying implications for future research.

# Chapter 2

# Visual salience

**Abstract**

Salient entities are assumed to be more accessible in memory, which makes them more likely to be referred to first and to be referred to with an attenuated expression, such as a pronoun. It is less clear, however, how different types of salience interact in influencing referent accessibility. In this chapter, we address the question whether non-linguistic factors can affect accessibility in the presence of a linguistic context. We present two story completion experiments in which we investigated the effect of visual salience (foregrounding) in interaction with linguistic salience (subjecthood) of two story characters both on the choice of referent and on the choice of referring expression. In Experiment 1, linguistic salience was moderated by inducing a topic shift in the discourse context. In Experiment 2, contexts in which linguistic salience was unclear were compared to contexts in which one of the characters was highly linguistically salient. The results show that visual salience influences referent choice independently of linguistic salience, but that it does not have an effect on the choice of referring expression. This suggests that visual salience has an influence on the global interpretation of the scene, but does not directly affect the accessibility status of individual entities. This is compatible with a view of language production in which utterance planning is influenced by conceptual and discourse factors rather than by low-level perceptual factors.

**This chapter is based on:**

Vogels, J., Krahmer, E. J., & Maes, A. A. (2013). Who is where referred to how, and why? The influence of visual saliency on referent accessibility in spoken language production. *Language and Cognitive Processes* 28 (9), 1323-1349.

## 2.1. Introduction

Reference is an important part of human communication. When we speak, we are constantly referring to objects or persons in our physical environment, to previous linguistic utterances, or to general knowledge about the world around us. A consequence of this is that our use of language is strongly interwoven with the context in which it is produced (Ariel, 1990). The ways in which context affects the production of references are not completely understood. Context may influence references in different parts of the production process. On the one hand, it may affect what people choose to refer to first, i.e., how they choose a 'starting point' for their utterance (MacWhinney, 1977). On the other hand, when a referent has been established, context may influence what type of referring expression people use to refer to it, ranging from elaborate descriptions involving full noun phrases and modifiers (e.g., *the blonde girl with the big earrings*) to short, low-informative elements such as pronouns (e.g., *she*). It is generally acknowledged that an important factor in guiding speakers both in what to refer to and in how to refer to it is salience. Salient entities are assumed to have a mental representation that is more activated, and hence are more readily available for the language production process (Levelt, 1989). Therefore, people tend to mention the most salient entity in the context first (e.g., Grosz, Joshi, & Weinstein, 1995; Osgood, 1971; Osgood & Bock, 1977; Stevenson, Crawley, & Kleinman, 1994). In addition, people tend to use more reduced expressions to refer to a salient entity, probably because highly accessible entities do not need an extensive description for identification (e.g., Ariel, 1990; Gundel, Hedberg, & Zacharski, 1993).

Different factors, both linguistic and non-linguistic ones, have been identified as contributors to salience. An entity is more salient, for example, when it is a subject (Gordon, Grosz, & Gilliom, 1993; Grosz et al., 1995), a topic (Ariel, 1990; Givón, 1983; Gundel et al., 1993), the source of an event (Arnold, 2001; Stevenson et al., 1994), animate (Prat-Sala & Branigan, 2000 and many others), large (Flores d'Arcais, 1975), or more imageable (Bock & Warren, 1985). However, it is not clear how different types of salience interact in making a referent more or less accessible, and which types of salience are relevant for which part of the production process. Especially in research on the choice of referring expressions, the focus has been on the influence of salience in the linguistic context. The role of non-linguistic factors, such as visual salience, is much less clear. Given the fact that people often refer to things that are present both in their physical environment and in the discourse context, an important question is how

visual salience affects reference in interaction with linguistic salience. In this chapter, we explore to what degree visual salience is used in reference, and how it interacts with the linguistic context of an utterance. We investigate both whether people are influenced by visual salience in choosing what to refer to first, and whether they are influenced by visual salience in their choice for a particular type of referring expression.

Research on the effect of the visual context on language production suggests that speakers use non-linguistic information in planning their utterances. For instance, visual attention has been found to affect syntactic structure. When people describe visual scenes in which visual attention is drawn to a particular figure, they are more likely to mention this figure first. When the figure is the patient of a transitive event, for example, people tend to use passives or predicates that take the perspective of the patient, such as 'flee' instead of 'chase' (Flores d'Arcais, 1975; Gleitman, January, Nappa, & Trueswell, 2007; Sridhar, 1988; Tomlin, 1997). These findings indicate that visually salient objects are more likely to be referred to first. It is not clear, however, how these effects of visual salience interact with other types of salience, such as the referent's discourse status. The studies cited above all made use of scene descriptions, in which no linguistic context was provided. Other studies have found that the salience of an entity in the linguistic context, for example whether it constitutes given information or whether it was a subject in the preceding sentence, affects the choice of referent for first mention (e.g., Bock & Irwin, 1980; Ferreira & Yoshita, 2003; Prat-Sala & Branigan, 2000; Stevenson, 2002). Referent choice may therefore depend on a combination of linguistic and non-linguistic salience, or one type of salience might override the other (cf. Prat-Sala & Branigan, 2000 for the interaction between discourse salience and animacy).

There is also evidence that visual information influences the choice of referring expression. For example, people use more elaborate expressions, such as modified noun phrases instead of bare noun phrases, when multiple possible referents are present that are visually ambiguous (Brown-Schmidt & Tanenhaus, 2006; Ferreira, Slevc, & Rogers, 2005; Sedivy, 2003). They also use more reduced expressions when the referent is visually in focus (Beun & Cremers, 1998), and fewer reduced expressions when another possible referent is visually present, even when there is no ambiguity (Fukumura, Van Gompel, & Pickering, 2010). These findings suggest that the more salient a referent is in the visual context, the more attenuated the referring expression will be.

Most studies addressing the influence of visual information on the choice of referring expression were again scene description tasks that lacked a linguistic context. However, the choice of referring expression is generally assumed to be greatly influenced by the referent's saliency in the preceding discourse context. According to theories of reference (e.g., Ariel, 1990; Givón, 1983; Grosz et al., 1995), pronouns are appropriate when they refer to the subject or the topic of the previous sentence. Longer, more informative expressions are used when the referent is not a subject or a topic. Besides grammatical function and topicality, linguistic factors that have been found to affect the choice of referring expression include recency, givenness, frequency of mention, thematic roles and syntactic position (e.g., Anderson, Garrod & Sanford, 1983; Ariel, 1990; Arnold, 1998; Clark & Sengul, 1979; Gernsbacher & Hargreaves, 1988; Givón, 1983; Gundel et al., 1993; Stevenson et al., 1994). Few studies have investigated effects of visual properties on the choice of referring expression when a linguistic context is present. This might be due to the assumption of some researchers (e.g., Ariel, 1998) that linguistic context is a far more important source of accessibility than the physical world. However, empirical evidence is typically lacking, the claim being mainly based on corpus studies of written texts (e.g., Ariel, 1998, 2001), a domain in which physical context generally does not play a large role.

Recently, story completion experiments have been used to examine pronoun use in references to a character in the presence of both a visual and a linguistic context. In a study by Arnold and Griffin (2007), participants had to describe a picture following the presentation of another picture in combination with a context sentence. The stimuli came in three versions: (1) a competitor character was present both visually and linguistically; (2) no competitor was present; (3) a competitor was present both visually and linguistically in the first scene but not visually present in the second. Target referent and competitor were referentially unambiguous (i.e., of different genders). Arnold and Griffin found that participants used fewer pronouns to refer to the target character in condition (1) than in condition (2), suggesting that the presence of a competitor decreased the target referent's accessibility. However, no difference in pronoun use was found between conditions (1) and (3), suggesting that it was the linguistic presence of the competitor in the context sentence that affected salience rather than its visual presence during sentence production.

In a similar experiment, Fukumura et al. (2010) did find that visual context influenced the choice of referring expression. In their study, participants were presented with two consecutive pictures showing either one toy character (competitor

not present condition) or two toy characters (competitor present condition). The first picture was combined with a written context sentence in which only the target character was mentioned (competitor not mentioned condition) or both characters were mentioned (competitor mentioned condition). In the second picture, the target character performed a simple action. Participants described this picture to a confederate, who then acted out the description using the real toys. The results showed that participants used fewer pronouns to refer to the target referent when the competitor was visually present than when it was not visually present. In addition, the effect of visual context was larger in the condition where the competitor was linguistically present than in the condition where the competitor was not mentioned at all. These findings suggest that visual information affects the referent's accessibility and therefore the likelihood that it will be referred to with a pronoun. They also suggest that the visual context has a greater effect when the linguistic context is less compelling (i.e., when the competitor is also linguistically present).

Still, in both Arnold and Griffin (2007) and Fukumura et al. (2010), as well as in other studies investigating the role of visual information on the production of referring expressions, visual information was manipulated differently than in the studies on referent choice cited above. The latter manipulated visual salience either by varying the objects' intrinsic perceptual properties, such as its size (Flores d'Arcais, 1975; Osgood, 1971; Sridhar, 1988), or by making use of (implicit) attentional cues, such as a black square presented very briefly in the same position as where the target figure appears immediately afterwards (Gleitman et al., 2007; Tomlin, 1997). The studies investigating choice of referring expression, on the other hand, varied the number of possible referents in the visual context, but not the intrinsic salience of the referents. With more than one possible referent, multiple entities have to be kept in memory at the same time. Because attention is spread over more than one entity, the individual activation of the entities may be reduced in the speaker's memory (cf. Arnold & Griffin, 2007). The influence of visual information on the choice of referring expression may thus be an effect of competition between possible referents, rather than an effect of salience.

To determine how salience affects the choice of referring expression, as well as the choice of referent, properties of the referent itself should be taken into account. Linguistically, referents that are topics or subjects are more salient than referents that are not topics or subjects (e.g., Arnold, 1998). Visually, properties such as size, centrality, color, foregrounding, orientation, intensity, and visual complexity have

been identified as important cues to salience (e.g., Clark & Chase, 1972; Coco & Keller, 2009; Henderson & Ferreira, 2004; Kelleher, Costello, & Van Genabith, 2005; Mazza, Turatto, & Umiltà, 2005; Parkhurst, Law, & Niebur, 2002). The more prominent these perceptual properties, the more attention they will receive by the visual system (Desimone & Duncan, 1995; Parkhurst et al., 2002). Since the role of these properties in determining the accessibility of a referent in discourse is still unclear, it remains an open question in what way visual salience affects the choice of referent and the choice of referring expression in interaction with linguistic context.

Another question is at which stage of the production process effects of visual salience take place. According to Griffin and Bock (2000), order of mention (e.g., agent before patient) is influenced by a global apprehension of the event (e.g., what the source and the goal of the action are), but not by the visual salience of individual elements in the scene. Thus, some higher-level conceptual knowledge is required before linguistic processing can take place. For example, a mental model of the situation may be created, including world knowledge and inferences, which guides the formulation of utterances (Anderson et al., 1983; Johnson-Laird, 1983; Morrow, 1985; Sanford & Garrod, 1981). Low-level perceptual factors will not directly affect utterance planning (Bock, Irwin, & Davidson, 2004; Griffin & Bock, 2000). By contrast, Gleitman et al. (2007) found that even in events inducing a conjoined NP subject, in which there is no clear instigator or recipient of the action, visual attention influenced order of mention. They took this as evidence that attention directly affects the activation of lemma representations, without the need for first apprehending the gist of the scene.

If visual attention indeed leads to a higher accessibility of lemma representations, which subsequently affects order of mention in an incremental manner (e.g., Kempen & Hoenkamp, 1987), one would expect that it also affects the choice of referring expression, since this requires access to the corresponding lemma (Levelt, 1989). If, on the other hand, visual attention only influences the global representation of the scene, it may still affect the choice of referent by determining from which perspective the scene is viewed. For example, the same event may be seen as a 'giving' event or as a 'receiving' event, depending on which entity is taken as the figure and which as the ground (Gleitman et al., 2007). This may influence which entity is mentioned first. However, according to this account, visual attention does not necessarily affect the choice of referring expression, since this is crucially dependent on the accessibility of

mental representations of the individual entities, and not on the representation of the scene as a whole.

We hypothesize that if referent accessibility is dependent on a combination of linguistic and non-linguistic factors, a referent's visual salience will interact with its linguistic salience in determining the accessibility of its representation in memory. In this case, an entity that is salient in the visual context may activate its corresponding mental representation(s) more than other entities, in the same way as a linguistically salient entity does. This may guide the speaker in choosing what to mention first, as well as in the choice of how to encode it in a referring expression. Thus, visually salient referents are expected to be chosen more often as first referents and to be more often referred to with reduced expressions than visually non-salient entities. In addition, if linguistic information is more important in determining accessibility than visual information, as suggested by previous studies, any effects of visual saliency should at least be expected in contexts where linguistic saliency is less clear.

Alternatively, linguistic and visual salience may affect referential choices in different ways. Visual salience may influence the global representation of the narrative, e.g., who the main characters are, but may not directly affect the accessibility of individual entities, as suggested by Griffin and Bock (2000). Under this hypothesis, no effect of visual salience on choice of referring expression is predicted, while there may still be an effect on referent choice via figure-ground assignment or protagonisthood.

To investigate these alternative hypotheses, we conducted two story completion experiments in Dutch, in which we manipulated visual salience as a property of the characters in the visual scene themselves, and in which linguistic salience in two context sentences was also varied. In the experiments, two characters were always present simultaneously, both visually and linguistically, but one character was made visually and/or linguistically salient relative to the other. To make the stories naturalistic, the visual scenes used were photographs of real people. We investigated the effects of a referent's visual salience on the likelihood that it is mentioned first and on the use of pronouns versus full noun phrases. In Experiment 1, linguistic saliency was moderated by inducing a topic shift between the two context sentences. In Experiment 2, linguistic saliency was moderated by employing a noun phrase conjunction in the first context sentence and a sentence conjunction in the second context sentence. This condition was compared to two control conditions in which the linguistic saliency manipulation was very strong. In addition, differences in visual

saliency were emphasized more than in Experiment 1. In both experiments, we found evidence for an effect of visual salience on referent choice, but not on choice of referring expression, suggesting that visual salience affects reference on a different level of processing than linguistic salience.

## 2.2. Experiment 1

### *2.2.1. Methods*

#### 2.2.1.1. Participants

Sixty-four undergraduate students (17 male, 47 female; aged 18 to 43; mean age 21.7) from Tilburg University participated in the experiment for course credit. All were native speakers of Dutch and had normal or corrected to normal vision.

#### 2.2.1.2. Materials

Twelve short stories were created that served as the experimental items. Each story consisted of two pictures, showing a male and a female character in a certain situation, accompanied by two sentences and the onset of a third sentence. The pictures and the sentences formed the context for the third sentence, which had to be completed by the participants. Sentences 1 and 2 accompanied the first picture of a pair, while the onset of the third sentence was aligned with the onset of the second picture. In the second picture, one character always performed an action (henceforth 'agent character'). The other character did not move with respect to the first picture. Which character performed the action was varied across items. In addition, the visual and the linguistic salience of the characters were manipulated. The four different picture pairs of an experimental item with the accompanying context sentences are exemplified in Figure 2.1.

Linguistic salience was manipulated by making one of the characters the subject of the first context sentence and the other one the subject of the second context sentence. The subject of the second sentence was considered to be linguistically salient, because it was the subject of the sentence directly preceding the sentence that had to be completed, and the most recent possible antecedent for a referring expression. This is in line with theories of reference, such as centering (Grosz et al., 1995), in which the subject of the previous utterance is considered the most salient entity. The subject of the first sentence was considered to be linguistically less salient, because referential

distance was longer (e.g., Ariel, 1990). Thus, one character was the subject of the second sentence (linguistically salient) and the other the subject of the first sentence (linguistically non-salient). The topic shift between the first and the second sentence was included to ensure that neither character became so linguistically salient that any effects of visual salience would be overruled.

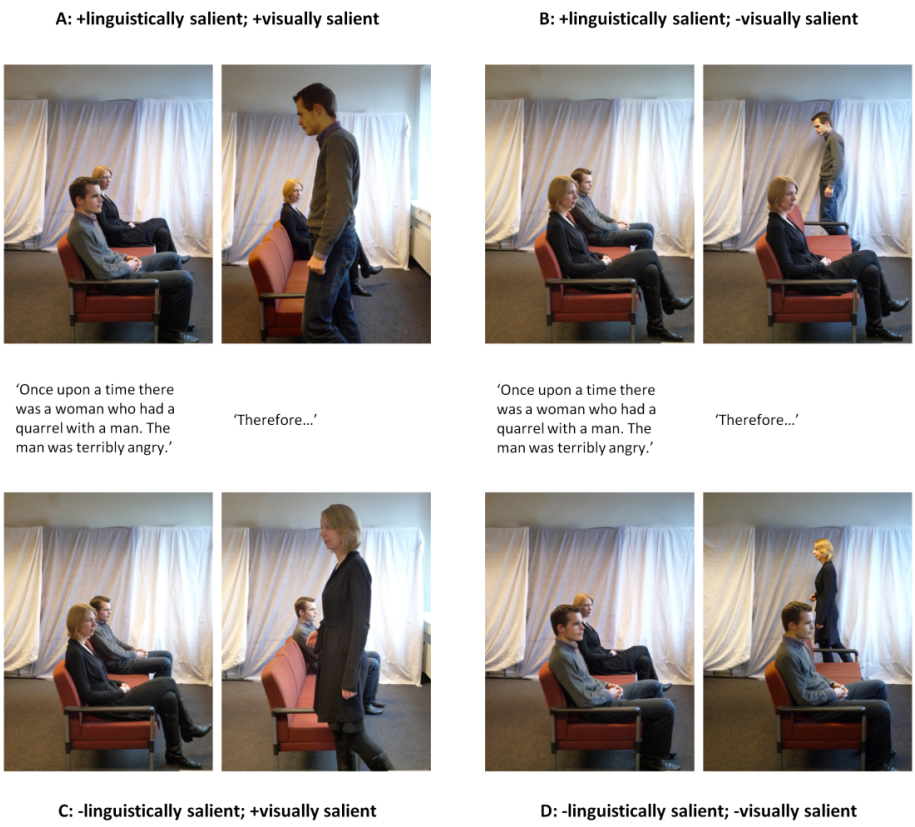


**Figure 2.1.** A stimulus item from Experiment 1 in four different conditions: (A) agent character (i.e., the person performing the action in the second picture) is both linguistically and visually salient; (B) agent character is linguistically but not visually salient; (C) agent character is visually but not linguistically salient; (D) agent character is neither linguistically nor visually salient. The corresponding context sentences are translations of the Dutch originals.

The context sentences all had the same structure, which is illustrated in Table 2.1. The first context sentence always started with the phrase *Er was eens* 'Once upon a time there was', followed by an indefinite subject, which referred to the female character (*een vrouw* 'a woman') in half of the cases and to the male character (*een man* 'a man') in the other half. The subject was modified by a relative clause describing the

situation (e.g., *die ruzie had* 'who had a quarrel'), always followed by a prepositional phrase introducing the other character (e.g., *met een man* 'with a man'). Subsequently, this character became the subject of the second sentence, which described a physical or emotional state (e.g., *De man was verschrikkelijk boos* 'The man was terribly angry'). Across items, the predicates in the first two sentences were varied, as well as the gender of the subject and PP constituents (i.e., the subject of sentence 1 referred to the female character and both the character in the PP and the subject of sentence 2 referred to the male character, or vice versa). The onset of the third sentence always consisted of the word *Daarom* 'Therefore'.[1] Because Dutch is a verb second language, this means that participants had to start their utterance with a finite verb, directly followed by the subject, which was the constituent of interest. All context sentences were recorded by a female native speaker of Dutch. The sentences were pronounced with a neutral statement intonation, and with no stress accents on the noun phrases mentioning the characters (e.g., main stress in the second context sentence was always on the state describing adjective).

**Table 2.1.** Template for the context sentences in Experiment 1.

| Sentence 1 | Sentence 2 | Onset of sentence 3 |
|---|---|---|
| *Er was eens {een vrouw, een man} die {PREDICATE} met {een man, een vrouw}.*[a] | *{De man, De vrouw} {was, had} {STATE}.* | *Daarom…* |
| 'Once upon a time there was {a woman, a man} who {PREDICATE} with {a man, a woman}.' | '{The man, The woman} {was, had} {STATE}.' | 'Therefore…' |

[a] Alternatives for constituents that vary across items are between curly brackets.

To test the consistency of the linguistic materials, we conducted a pretest in which participants provided written completions of the sentence combinations (without the pictures). The results showed no biases in particular items for the use of certain types of referring expression (One Way ANOVA with multiple comparisons: all *p*

[1] The connective *Daarom* 'Therefore' was chosen to ensure a close connection between the two context sentences, as well as between the two pictures. Because research has shown that the type of connective affects the choice of referent for next mention (e.g., Stevenson et al., 1994), care was taken that continuations with either character resulted in plausible stories, given the previous context sentences (see below).

values > .1). A second pretest in which participants had to choose between either a pronoun or an NP for a fixed referent also showed no significant differences between the items (all *p* values > .5).

Visual salience in the pictures was manipulated by having one of the characters appear in the foreground and in a central position in the picture, while the other one appeared in the background in a more peripheral position. In most cases the foregrounded character also partly occluded the backgrounded character. Varying both the visual salience of the two characters and the agent character in the second picture resulted in four versions of each picture pair (either the male or the female character was visually salient; either the male or the female character was the agent; see Figure 2.1). The pictures were photographs taken with a digital camera. Two couples posed for all pictures, which were all taken in the same room, with a neutral background. A statistical test of the size of both characters (size in pixels, measured from head to foot) confirmed that across pictures, the characters that were intended to be visually salient were depicted significantly larger than the characters that were intended to be non-salient ($t_{142}$ = 19.375; $p$ < .001). To avoid any effects of the left-to-right orientation of the characters in the pictures, mirror versions were created for all picture pairs (not shown in Figure 2.1). In all, each story had eight different versions.

In the first picture of each story, both characters were in a neutral position (e.g., sitting next to each other on the couch). In the second picture, either the male character performed a simple action (e.g., walking away), or the female character performed the same action. Care was taken that the third sentence could be finished in both versions of the second picture, i.e., when the man was the agent and when the woman was the agent. Therefore, the action depicted in the second picture had to be compatible with the preceding context in both versions. For example, the action of walking away in reaction to the man being angry can be performed by both characters, since this is a plausible reaction for people both when they are angry themselves and when they are faced with anger from someone else.

In addition to the 12 experimental items, 16 filler items and 3 practice items were constructed. The fillers were identical to the experimental items, except for the fact that more variation was included in the situations: 5 items had one male and one female character, in 5 items there was only one character and in another 6 items there were two characters of the same gender. In addition, the characters were not always referred to as ‘a man’ and ‘a woman’, but they also sometimes had roles like ‘a teacher’ or ‘a salesman’. There was only one version of each filler and practice item. The items

were distributed over four lists according to a Latin Square design, such that each list contained one condition of a stimulus item. On each list, half of the items were in the mirrored version, and another four lists were created in which the other half of the items were mirrored, thus resulting in a total of eight lists. For each list, items were quasi-randomized, with the filler items in a fixed position and two experimental items never occurring in consecutive slots.

#### 2.2.1.3. Procedure

Participants sat in a low noise cabin in front of a computer screen. They had a keyboard at their disposal, which was only used to start the experiment. Between the keyboard and the computer screen was a microphone. The experiment was assembled and run with the E-Prime 2.0 software program. Participants were instructed to complete each story initiated by the two context sentences in such a way that it would fit with the situation shown in the second picture. They were not otherwise instructed about the content of their responses: Participants were free to complete the stories in any way they liked, with the only restriction that their first sentence had to be connected to the word *Daarom* 'Therefore'. They were not allowed to repeat this word, because this would cause a break in the continuation of the story. Participants were further instructed to use their first intuitions about how to complete the story and not to ponder too long. The first three trials were practice trials, after which participants had the opportunity to ask any remaining questions.

Pressing the space bar started the experiment. First, the trial number appeared on the screen for 1500 ms, accompanied by a 500 ms beep. Next, a fixation cross was shown for 600 ms, after which the first picture appeared. Immediately with the first picture, the first two sentences of the story were presented over the computer speakers. The second picture was presented 700 ms after termination of the second sentence, together with the word *Daarom*. Recording started at the same time. As soon as the word *Daarom* had sounded, there was an 8 s pause in which the second picture remained on the screen and the participant could complete the sentence. When the 8 s had elapsed, recording stopped and the next trial was started. It took about 10 minutes to complete the experiment.

#### 2.2.1.4. Data coding

After discarding the filler and practice items, the remaining (12 x 64 =) 768 recordings of the participants' story completions were scored for two variables: (1) which

character participants referred to as the subject of the sentence that started with *Daarom* 'Therefore', and (2) which type of referring expression participants used for this referent. For the choice of referent, we coded references to the agent character as 'agent' and references to the non-agent character as 'non-agent'. For the type of referring expression, the following codings were employed: NPs preceded by a definite article (*de man* 'the man'), a demonstrative (*die vrouw* 'that woman') or an adjective (*de boze man* 'the angry man') were coded as 'NP'; full pronouns (*hij* 'he', *zij* 'she') and reduced pronouns (*ie, die* 'he', *ze* 'she') were coded as 'pronoun'. Other types of referring expression were not attested.

All first references after the word *Daarom* 'Therefore' occurring as the subject of the completion sentence were scored. All other references were ignored. We excluded 23 responses from the data set in which reference was made to both characters at the same time, and 18 in which neither character was mentioned as the subject. In addition, we removed 10 cases in which the word *Daarom* was repeated, 5 cases in which the referring expression used was not clear, 4 cases in which the response suggested a misinterpretation of the story with respect to the pictures, and 2 cases in which the recording did not contain any speech. In all, 62 trials (8%) were excluded from the data set, resulting in 706 useful trials for the analysis.

2.2.1.5. Design and statistical analyses

Crossing the two independent variables resulted in a 2 (agent character is linguistically salient or linguistically non-salient) x 2 (agent character is visually salient or visually non-salient) within-subjects and within-items design. The dependent variables were the proportion of subject references to the agent character out of all references and the proportion of pronoun references to the character mentioned as the subject out of all subject references to that character. For the latter variable, references were analyzed separately for each character type (agent or non-agent). However, because there were too few data points for the non-agent references to perform statistical analyses, only the results for the references to the agent character will be reported. We conducted two logit mixed model analyses (Jaeger, 2008), using the lme4 package in the R software program (www.r-project.org): one over the log odds of a subject reference to the agent character, and one over the log odds of a pronoun reference to the agent character. In both cases, linguistic and visual salience of the agent character were included as fixed factors, and participants and items as random factors. The fixed factors were centered to reduce collinearity between

predictors. Random intercepts and random slopes for participants and items were included to account for between-subject and between-item variation. Starting with a model with a full random effect structure, we used model comparisons to determine whether the inclusion of a random slope was justified by the data. Random slopes that did not contribute to the fit of the model according to a likelihood ratio test were removed (Jaeger, 2011). Only the final models will be reported.

### *2.2.2. Results*

#### 2.2.2.1. Choice of referent

In the majority of the cases (624, 88%), participants referred to the agent character as the subject of their response. Still, there were 82 responses (12%) in which participants referred to the non-agent character (i.e., the character that did not move with respect to the first picture) in subject position. There was no difference between the mirrored and the unmirrored versions of the pictures, i.e., participants were not more likely to refer to the agent character when it was on the left than when it was on the right, $\chi^2(1) < 1$, $p = .98$.

Figure 2.2 shows the proportion of subject references to the agent character as a function of its linguistic and visual salience. The final logit mixed model is summarized in Table 2.2. Firstly, we found an effect of linguistic salience on referent choice: Surprisingly, subject references to the agent character were more frequent (95.5%) when this character was *not* linguistically salient (i.e., when it was the subject of the first context sentence), than when it was linguistically salient (i.e., when it was the subject of the second context sentence) (81%). This difference was significant (see Table 2.2). The negative coefficient in Table 2.2 means that the predicted odds for a subject reference to a linguistically salient agent character are lower than the odds for a subject reference to a linguistically non-salient agent character.

Secondly, we found an effect of visual salience on referent choice: Subject references to the agent character were more frequent (91.1%) when this character was visually salient, than when it was visually non-salient (85.5%). This difference was significant (see Table 2.2). We found no significant interaction between linguistic and visual salience. By-subjects and by-items random slopes for linguistic salience were included, as they significantly improved model fit, $\chi^2(2) = 8.00$, $p < .05$ and $\chi^2(2) = 42.33$, $p < .001$, respectively. This indicates that participants and items varied in the degree to which they were affected by linguistic salience.

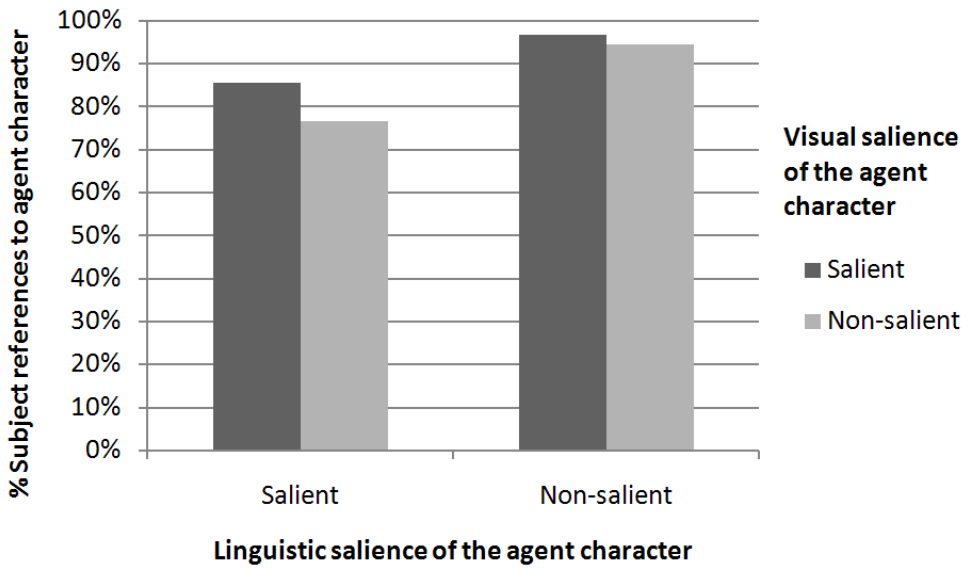


**Figure 2.2.** Percentage of subject references to the agent character out of all references in Experiment 1, plotted against its linguistic and visual salience.

**Table 2.2.** Logit mixed model for referent choice in Experiment 1 (*Ling. sal.* = Linguistic salience; *Vis. sal.* = Visual salience).

| Random effects | | $s^2$ | | |
|---|---|---|---|---|
| *Subjects* | *Intercept* | 1.45 | | |
| | *Ling. sal.* | 4.28 | | |
| *Items* | *Intercept* | 0.57 | | |
| | *Ling. sal.* | 9.49 | | |
| Fixed effects | $\beta$ | *SE* | *z* | *p* |
| *Intercept* | 3.81 | 0.39 | 9.85 | < .001 |
| *Ling. sal.* | -2.28 | 1.06 | -2.15 | < .05 |
| *Vis. sal.* | 0.82 | 0.41 | 1.99 | < .05 |
| *Ling. sal. * Vis. sal.* | 0.15 | 0.81 | 0.19 | .85 |

2.2.2.2. Choice of referring expression

For all subject references to the agent character (n = 624), we determined whether participants used a pronoun or a full noun phrase. The results for the effects of linguistic and visual salience on pronoun use are presented in Figure 2.3. The final logit mixed model is summarized in Table 2.3.

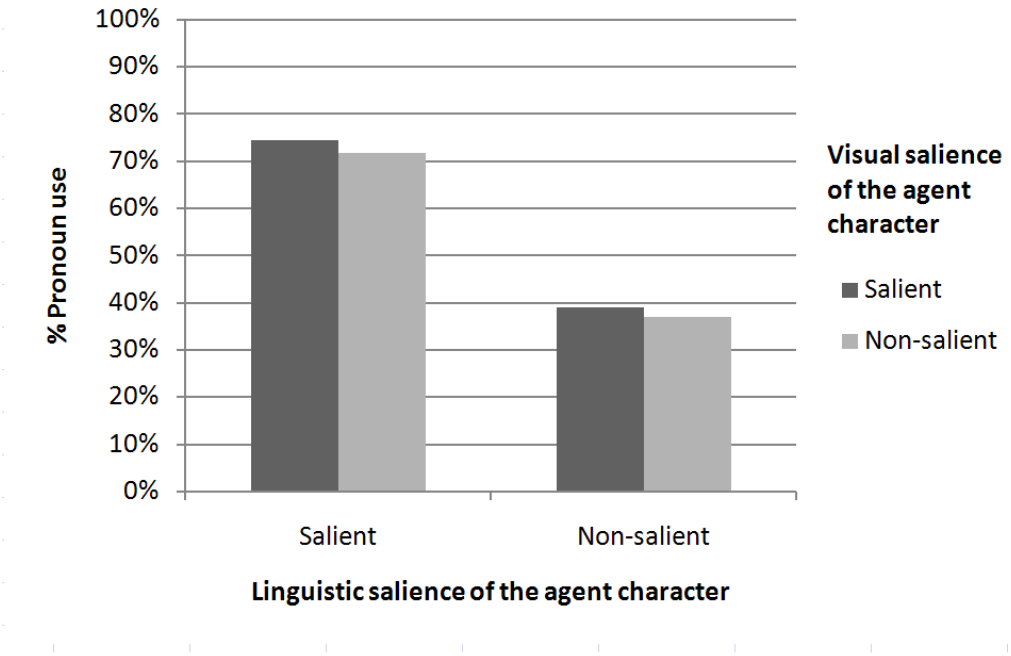


**Figure 2.3.** Percentage of pronoun references to the agent character out of all references (pronoun and full NP) to the agent character in Experiment 1, plotted against its linguistic and visual salience.

**Table 2.3.** Logit mixed model for choice of referring expression in Experiment 1 (*Ling. sal.* = Linguistic salience; *Vis. sal.* = Visual salience).

| Random effects | | $s^2$ | | |
|---|---|---|---|---|
| *Subjects* | *Intercept* | 2.31 | | |
| *Items* | *Intercept* | 0.70 | | |
| Fixed effects | $\beta$ | *SE* | *z* | *p* |
| *Intercept* | 0.36 | 0.33 | 1.09 | .27 |
| *Ling. sal.* | 2.25 | 0.23 | 9.93 | < .001 |
| *Vis. sal.* | 0.16 | 0.21 | 0.78 | .43 |
| *Ling. sal.* * *Vis. sal.* | 0.08 | 0.42 | 0.18 | .86 |

We found a main effect of linguistic salience: Pronouns were used more frequently when the agent character was linguistically salient (73%) than when it was linguistically non-salient (38%). This difference was significant (see Table 2.3). We found no main effect of visual salience (57% vs. 54% pronoun use), and no interaction

between linguistic salience and visual salience. By-subjects and by-items random slopes were not included, as they did not improve model fit.

*2.2.3. Discussion*

Experiment 1 showed effects of visual salience on the choice of referent, but not on the choice of referring expression. Participants used the visual salience of an entity to determine whether they would refer to it first (as the subject), but not to determine whether they would refer to it with a pronoun or a full NP. This suggests that visual salience does not affect the accessibility of mental representations associated with the individual characters in the scene, but is used for the global interpretation of the story, which subsequently influences figure-ground assignment or protagonisthood. In addition, there were no interactions with linguistic salience, indicating that the effects of visual salience were not influenced by whether the referent was the subject of the first or the second context sentence. This suggests that visual salience and linguistic salience affect reference production independently.

There may be two problems with the findings of Experiment 1. Firstly, the effect of linguistic salience on the choice of referent was not as predicted. The subject of the second context sentence, being the most recent subject, was expected to be more salient and hence to induce more subject references in the participants' response. However, participants were more likely to refer to the subject of the first context sentence than to that of the second. An example is given in (1c), (1a) and (1b) being the preceding context sentences (the participant's utterance is presented in bold).

(1) a. *Er was eens een man die stond te kletsen met een vrouw.*
'Once upon a time there was a man who was chatting with a woman.'

b. *De vrouw was een beetje duizelig.*
'The woman was a bit dizzy.'

c. *Daarom* ***zei hij***$_{\text{NON-AGENT}}$ ***dat ze***$_{\text{AGENT}}$ ***een stoel moest pakken zodat ze erop kon gaan zitten.***
'Therefore he$_{\text{NON-AGENT}}$ said that she$_{\text{AGENT}}$ should take a chair so that she could sit down on it.'

In (1c), the subject pronoun *hij* 'he' refers to the subject *een man* 'a man' in (1a), even though in (1b) the subject has shifted to *de vrouw* 'the woman', and even though the referent is not the agent. This illustrates participants' preference to continue the story

with the subject of the first context sentence, rather than with that of the second. We propose that this preference results from an interpretation of the subject of the first context sentence as the discourse topic or the protagonist of the story. The protagonist is the character from whose perspective the story unfolds. The storyteller takes the viewpoint of the protagonist, and expresses this by putting it in subject or topic position right at the beginning of the discourse (Morrow, 1985). In the present experiment, the interpretation of the subject of the first sentence as the protagonist might have been encouraged by the fact that it was introduced with the words *Er was eens* 'Once upon a time there was'. This may have caused the participants to continue referring to this character as the subject of their utterances (cf. Anderson et al., 1983).

This effect of protagonist assignment on referent choice may have had consequences for the effect of visual salience. Visual salience may also be a factor that affects the chance that a character is seen as the main character (or the 'figure') of the story. If linguistic and visual salience are indeed additive in determining the protagonist and therefore the choice of referent, this protagonisthood effect should be largest when the two factors are congruent, i.e., when the referent is either both the subject of the first context sentence and visually salient, or both the subject of the second context sentence and visually non-salient. Figure 2.2 indeed suggests that this is the case. In addition, although we did not find a significant interaction between linguistic and visual salience, Figure 2.2 suggests that the effect of visual salience tends to be larger when salience is otherwise unclear (e.g., the referent is an agent but not linguistically introduced as the protagonist). Nevertheless, to single out the effect of visual salience on referent choice, linguistic contexts that are more neutral with respect to protagonist assignment should be investigated.

A second problem is that the fact that we did not find effects of visual salience on the type of referring expression could have been due to the dominance of linguistic salience: Information from the linguistic context could have been already sufficient to choose an expression, obscuring any effects of visual salience. It might be the case that visual salience only affects the choice of referring expression when a referent's salience in the linguistic context is unclear. It also might be the case that our manipulation of visual salience in Experiment 1 was not strong enough, or too heterogeneous across stimulus items to be manifest in the story continuations. In addition, the stimulus items may have differed in the degree of coherence between the context sentences. This might have resulted in an interpretation of the second picture

as unconnected to the preceding context. Experiment 2 was set up to deal with these issues.

## 2.3. Experiment 2

Experiment 2 was similar to Experiment 1, but a number of adaptations were employed to clarify the precise interplay between linguistic and visual salience in the choice of referent and the choice of referring expression. Firstly, we constructed context sentences in which both characters were expected to be about equally salient. These contexts were compared to contexts in which one character was made highly salient with respect to the other. If visual salience affects the choice of referring expression when linguistic salience is indecisive, an effect of visual salience is predicted in the former contexts, but not in the latter. Secondly, some adjustments to both the visual and the linguistic context were made to reduce differences in salience across items, to increase coherence between the context sentences, and to boost effects of visual salience.

### *2.3.1. Method*

#### 2.3.1.1. Participants

Forty-eight students (14 male; 34 female; aged 18 to 52; mean age 22.2) from Tilburg University participated for course credit. They were all native speakers of Dutch and had normal or corrected to normal vision. None of them had participated in Experiment 1.

#### 2.3.1.2. Materials

Twelve new short stories were created, similar to the stimuli in Experiment 1. As before, they involved two characters, whose visual and linguistic salience were manipulated. Visual salience had two levels (salient or non-salient), and linguistic salience had three levels (salient, non-salient, or undetermined). This resulted in six versions of each stimulus item. An example of an experimental item is shown in Figure 2.4.


**Figure 2.4.** A stimulus item from Experiment 2 in six different conditions: (A) agent character is both linguistically and visually salient; (B) agent character is linguistically but not visually salient; (C) agent character is visually but not linguistically salient; (D) agent character is neither linguistically nor visually salient; (E) agent character is visually salient and neither linguistically salient nor linguistically non-salient; (F) agent character is visually non-salient and neither linguistically salient nor linguistically non-salient. The corresponding context sentences are translations of the Dutch originals.

In the linguistically salient conditions (condition A & B in Figure 2.4), the agent character was the subject of both preceding context sentences. In the linguistically non-salient conditions (condition C & D in Figure 2.4), it was only mentioned in a prepositional phrase in the first context sentence. In the undetermined conditions (condition E & F in Figure 2.4), the second context sentence was a coordinated sentence, of which the first member had the agent character as the subject. The other character was mentioned as the subject of the second member. In the first context sentence, both characters were mentioned as the subject in a coordinated NP. This structure was chosen because it was assumed on the basis of the centering framework (Grosz et al., 1995) that when both characters are subjects in the directly preceding linguistic context, this makes their accessibility status unclear.

The sentences further differed from those in Experiment 1 in the following ways (see also Figure 2.4): Firstly, the first context sentence no longer started with *Er was eens…* 'Once upon a time there was…', to avoid a strong linguistic cue for protagonist assignment. Secondly, instead of being exclusively referred to as *man* 'man' and *vrouw* 'woman' in the experimental items, the characters were called *jongen* 'boy' and *meisje* 'girl' in half of the items, to increase variation in the descriptions for the characters. In the second context sentence, the adjective describing a physical or emotional state of the character(s) was always stage-level, i.e., it denoted a temporary, event-like property, such as *geïrriteerd* 'irritated' or *moe* 'tired' (Carlson, 1977). In contrast to individual-level predicates, which describe more or less permanent states (e.g., 'arrogant' or 'honorable'), stage-level adjectives are more likely to induce an episodic reading of the story, and make it less likely that the second picture will be described as a habitual or generic event. In the linguistically undetermined conditions (E & F), the second sentence described a physical or emotional state of both characters, e.g., *Het meisje was vol goede wil maar de jongen raakte enorm gepikeerd* 'The girl was full of good will but the boy got really annoyed'. In this way, both characters had the same thematic role (Experiencer), which should prohibit any differences in salience arising from this factor (see, e.g., Stevenson et al., 1994). Finally, to further emphasize the episodic nature of the stories, the finite verb in the second sentence was changed from static *zijn* 'to be' to dynamic *worden* 'to become' or *raken/krijgen* 'to get' (e.g., *het meisje werd tamelijk moe* 'the girl became pretty tired'). All sentences were recorded by the same speaker as in Experiment 1. Care was taken that in condition E and F the coordinated clauses in the second context sentence formed a single prosodic unit, such that they would not be interpreted as separate sentences.

Visual salience was manipulated in the same way as in Experiment 1: The agent character in the second picture was either foregrounded (visually salient) or backgrounded (visually non-salient). However, some additional photographic means were employed to emphasize the difference in visual salience. Firstly, a spotlight was put on the character in the foreground, making that person appear brighter than the visually non-salient character. Secondly, the character in the background was blurred a little by putting the camera's focus on the character in the front. Next, measures were taken to avoid large differences in the manipulation of visual salience between items. First of all, the positions of the two characters were kept constant across items, such that the distance between the salient and the non-salient character was always the same. Secondly, the action in the second picture always involved at least standing up from a chair, causing the agent character to be upright at the onset of the third sentence. Furthermore, the only pieces of furniture used were two chairs and an optional table, and photographing was done against a white screen. This was done to minimize distraction from the two characters caused by other objects. Finally, the actions performed in the second picture of a pair were made more similar in appearance, each being one of two kinds: either getting an object related to the state of the character described in the second sentence (e.g., getting a pillow when tired, either for oneself or for the other character), or movements, in particular leaving/walking away (e.g., after either the agent or the other character got irritated).

In addition to the 12 experimental items, 16 filler items and 4 practice items were constructed. These were similar to the filler items of Experiment 1. The items were distributed over a total of six stimulus lists. As in Experiment 1, mirror versions of all pictures were created. Since the left-to-right orientation of the characters in the pictures had not shown an effect on referent choice or choice of referring expression in Experiment 1, mirroring was done between items, instead of presenting each item in both a mirrored and an unmirrored version across lists. Items were distributed quasi-randomly over each list.

#### 2.3.1.3. Procedure

The procedure was similar to that of Experiment 1. After the participant had read the instructions, a practice block was started consisting of 4 practice items. Before the real experiment started, participants had the opportunity to ask any remaining questions. The setup of the experiment was the same as in Experiment 1. It took participants about 10 minutes to complete the experiment.

2.3.1.4. Data coding

After discarding the filler and practice items, (12 x 48 =) 576 responses to the experimental items remained. These were scored for referent choice and type of referring expression. The same coding scheme as in Experiment 1 was used. We excluded 1 response in which reference was made to both characters at the same time, and 1 response in which neither character was mentioned. We also excluded 1 response in which the word *Daarom* 'Therefore' was repeated, and another response of which the syntax did not match the V2 structure initiated by *Daarom*. Finally, we excluded 3 cases in which the referring expression was unclear or missing. In all, 7 responses (1.2%) were excluded, equally spread over the conditions, leaving 569 responses for analysis.

2.3.1.5. Design and statistical analyses

Crossing the two independent variables resulted in a 3 (agent character is linguistically salient, agent character is linguistically non-salient or linguistic salience is undetermined) x 2 (agent character is visually salient or visually non-salient) within-subjects and within-items design. The dependent variables were the proportion of subject references to the agent character out of all references, and the proportion of pronoun references to the character mentioned as the subject out of all subject references to that character. As in Experiment 1, only the results for references to the agent character will be reported for the latter variable. Statistical analyses were the same as in Experiment 1. Because linguistic salience was now a predictor with three levels, it was recoded into two binary factors using contrast (sum) coding. The first factor represents the difference between linguistically non-salient characters and characters whose linguistic salience is undetermined. The second factor represents the difference between linguistically salient characters and characters whose linguistic salience is undetermined.

*2.3.2. Results*

2.3.2.1. Choice of referent

In 544 cases (96%), participants referred to the agent character as the subject of their utterance. In 25 cases (4%), the subject referred to the non-agent character. The results for the proportion of subject references to the agent character as a function of its linguistic and visual salience are presented in Figure 2.5. The final logit mixed model is summarized in Table 2.4.

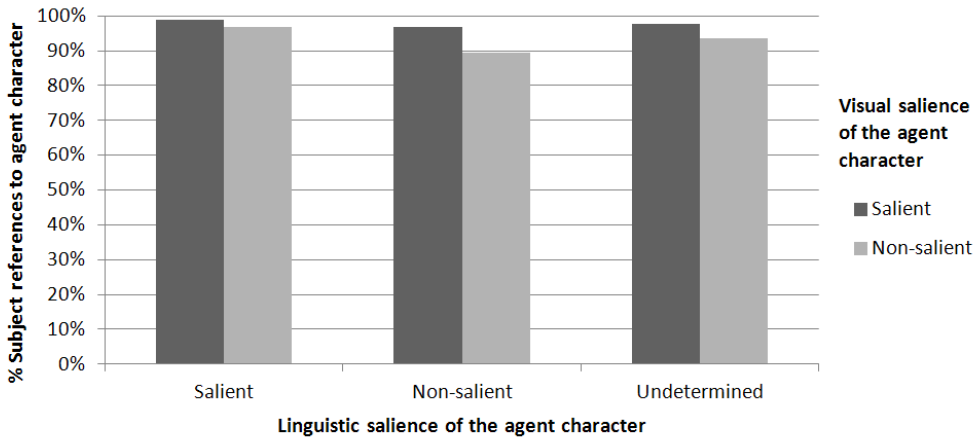


**Figure 2.5.** Percentage of subject references to the agent character out of all references in Experiment 2, plotted against its linguistic and visual salience.

**Table 2.4.** Logit mixed model for referent choice in Experiment 2 (*Ling. sal.* = Linguistic salience; *Vis. sal.* = Visual salience; *Sal.* = Salient; *Non-sal.* = Non-salient; *Undet.* = Undetermined).

| Random effects | | $s^2$ | | |
|---|---|---|---|---|
| *Subjects* | *Intercept* | 11.77 | | |
| | *Ling. sal.: Non-sal. vs. Undet.* | 46.49 | | |
| | *Ling. sal.: Sal. vs. Undet.* | 74.74 | | |
| *Items* | *Intercept* | 1.85 | | |
| Fixed effects | $\beta$ | *SE* | *z* | *p* |
| *Intercept* | 9.33 | 1.47 | 6.33 | < .001 |
| *Ling. sal.: Non-sal. vs. Undet.* | -0.16 | 2.43 | -0.07 | .95 |
| *Ling. sal.: Sal. vs. Undet.* | 0.78 | 3.60 | 0.22 | .83 |
| *Vis. sal.* | 2.25 | 0.84 | 2.67 | < .01 |
| *Ling. sal.: Non-sal. vs. Undet. * Vis. sal.* | 0.23 | 1.77 | 0.13 | .90 |
| *Ling. sal.: Sal. vs. Undet * Vis. sal.* | -0.68 | 2.43 | -0.28 | .78 |

Firstly, we found no effects of linguistic salience on referent choice: Although subject references were slightly more frequent for linguistically salient agent characters (i.e., being the subject of both context sentences; 98%), than for linguistically undetermined agent characters (i.e., both characters were subjects in both context

sentences; 96%), and slightly less frequent for linguistically non-salient agent characters (i.e., being only present in a PP in the first sentence; 93%), these differences were not significant (see Table 2.4). A Tukey's test of multiple comparisons showed that the differences between all three levels were non-significant, $\beta = 0.94$, SE = 3.65, $z = 0.26$, $p = .96$ for the difference between linguistically salient and linguistically non-salient agent characters; $\beta = 0.16$, SE = 2.41, $z = 0.07$, $p = 1$ for the difference between linguistically undetermined and linguistically non-salient agent characters; and $\beta = -0.78$, SE = 3.58, $z = -0.22$, $p = .97$ for the difference between linguistically undetermined and linguistically salient agent characters. Using model comparisons, the overall main fixed effect of linguistic salience also turned out to be non-significant, $\chi^2(4) = 0.48$, $p = .98$.

Secondly, we found an effect of visual salience on referent choice: References to the agent character were more frequent when it was visually salient (98%) than when it was visually non-salient (93%). This difference was significant (see Table 2.4). We found no significant interactions between linguistic and visual salience. By-subjects random slopes for linguistic salience were included, as they significantly improved model fit, $\chi^2(5) = 13.59$, $p < .05$, indicating that participants differed in the way they were affected by linguistic salience.

2.3.2.2. Choice of referring expression

For all subject references to the agent character (n = 544), we determined whether participants used a pronoun or a full noun phrase. The results for the effects of linguistic and visual salience on pronoun use are presented in Figure 2.6. The final logit mixed model is summarized in Table 2.5.

Firstly, we found an effect of linguistic salience: Pronouns were used more frequently (24%) when the linguistic salience of the agent character was undetermined than when the agent character was linguistically non-salient (11%), but less frequently than when it was linguistically salient (88%). The difference between linguistically non-salient and linguistically undetermined agent characters was significant, as was the difference between linguistically salient and linguistically undetermined agent characters, (see Table 2.5). A Tukey's test of multiple comparisons showed that the differences between all three levels were significant, $\beta = 16.57$, SE = 2.35, $z = 7.04$, $p = < .001$ for the difference between linguistically salient and linguistically non-salient agent characters; $\beta = 3.43$, SE = 0.93, $z = 3.70$, $p < .001$ for the difference between linguistically undetermined and linguistically non-salient agent characters; and $\beta = -$

13.14, SE = 2.27, $z$ = -5.79, $p$ < .001 for the difference between linguistically undetermined and linguistically salient agent characters. Using model comparisons, the overall main fixed effect of linguistic salience was found to be significant, $\chi^2(4)$ = 408.27, $p$ < .001.

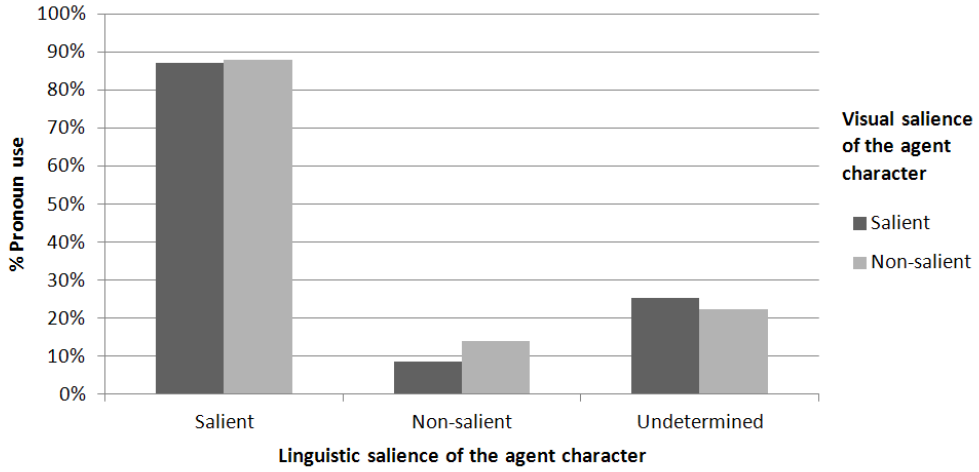


**Figure 2.6.** Percentage of pronoun references to the agent character out of all references (pronoun and full NP) to the agent character in Experiment 2, plotted against its linguistic and visual salience.

**Table 2.5.** Logit mixed model for choice of referring expression in Experiment 2 (*Ling. sal.* = Linguistic salience; *Vis. sal.* = Visual salience; *Sal.* = Salient; *Non-sal.* = Non-salient; *Undet.* = Undetermined).

| Random effects | | $s^2$ | | |
|---|---|---|---|---|
| *Subjects* | *Intercept* | 50.16 | | |
| | *Vis. sal.* | 21.83 | | |
| *Items* | *Intercept* | 1.32 | | |
| Fixed effects | $\beta$ | *SE* | *z* | *p* |
| *Intercept* | -2.60 | 1.22 | -2.12 | < .05 |
| *Ling. sal.: Non-sal. vs. Undet.* | -3.46 | 0.92 | -3.75 | < .001 |
| *Ling. sal.: Sal. vs. Undet.* | 13.25 | 2.30 | 5.77 | < .001 |
| *Vis. sal.* | -1.70 | 1.01 | -1.69 | .09 |
| *Ling. sal.: Non-sal. vs. Undet. * Vis. sal.* | -3.33 | 1.74 | -1.91 | .06 |
| *Ling. sal.: Sal. vs. Undet * Vis. sal.* | 3.01 | 2.18 | 1.39 | .17 |

Secondly, we found only a marginally significant effect of visual salience (see Table 2.5), showing a small tendency for fewer pronouns for visually salient agent

characters. There was also a marginally significant interaction between linguistic salience and visual salience (see Table 2.5), suggesting that the trend for fewer pronouns for visually salient referents only holds for linguistically non-salient referents. The difference in the effect of visual salience between linguistically salient and linguistically undetermined referents was non-significant. Recoding the predictor for linguistic salience revealed that the difference in the effect of visual salience between linguistically salient and linguistically non-salient referents was significant, $\beta = 6.49$, SE = 2.37, $z = 2.74$, $p < .01$: For linguistically salient referents, being visually salient results in more pronoun use than for linguistically non-salient referents. However, model comparisons showed that the overall interaction between linguistic and visual salience was only marginally significant, $\chi^2(2) = 4.68$, $p = .096$.

A by-subjects random slope for visual salience was included, as it significantly improved model fit, $\chi^2(2) = 9.89$, $p < .01$. This suggests that participants varied in the way they were affected by the manipulation of visual salience. Without the random slope, the main effect of visual salience and the interaction between linguistic and visual salience lost their (marginal) significance, which may suggest that these effects do not generalize to all participants.

*2.3.3. Discussion*

Experiment 2 was conducted to investigate whether a character's visual salience influences referent choice and the choice of referring expression when the character's salience in the linguistic context remains unclear. In such a case, the information from the linguistic context might be insufficient for a speaker to experience a strong preference to continue the story with one of the characters, or to choose a particular type of referring expression. The salience of the character in the visual context might then become decisive in making these choices. In Experiment 2, we attempted to create such a situation by on the one hand including a condition in which both characters were salient (i.e., subjects) throughout the linguistic context, and on the other hand strengthening the manipulation of visual salience.

The results show that, as in Experiment 1, visual salience affects the choice of referent, and does so independently of linguistic salience. Conversely, linguistic salience strongly affects the choice of referring expression, as expected, but it does not influence the choice of referent, which is different from what we found in Experiment 1. This may be related to the difference in the manipulation of linguistic salience between the two experiments (i.e., whether or not one of the characters was

introduced as the protagonist by the clause *Er was eens…* 'Once upon a time there was…'). The results of Experiment 2 further reveal a small effect of visual salience on the choice of referring expression that we did not find in Experiment 1: Referents that are visually salient are slightly *less* likely to be referred to with a pronoun. The interaction effects suggest that this effect is mainly carried by references to linguistically non-salient referents (see also Figure 2.6). Thus, for linguistically non-salient referents, which are already unlikely to be referred to with attenuated expressions, being visually salient reduces the chances of a pronoun even more. One possible explanation for this is that visual salience matches with linguistic focus: In the linguistically non-salient conditions (C & D), the referent is linguistically in focus in the first context sentence. Focus is associated with high informational importance and new information. Visual salience might convey similar properties. Thus, the tendency to use full NPs to refer to a linguistically focused referent might be encouraged when this referent is visually salient.

However, the fact that the effects of visual salience disappear when the by-subjects random slope for visual salience is removed suggests that they may be due to deviant behavior of some participants. Indeed, close inspection of the data revealed two participants that used full NPs only for linguistically non-salient referents that were visually salient. Without these two participants, the percentage of pronoun use in the linguistically non-salient conditions is about equal between the two visual conditions (around 9%). For linguistically salient referents and for referents whose linguistic salience is undetermined, visual salience does not seem to affect pronoun use, replicating our findings from Experiment 1. Thus, Experiment 2 shows that even when the referent's salience from the linguistic context is unclear, visually salient referents are not referred to with more attenuated expressions. This suggests that the lack of an effect of visual salience on the choice of referring expression in Experiment 1 is not due to the dominance of the linguistic context.

These results again suggest that the decision which character to mention first (as the subject) and the choice of referring expression are two processes that are affected by different kinds of information. Firstly, our finding that in the linguistically undetermined conditions (E & F) visual salience affects the choice of referent but does not affect the choice of referring expression suggests that visual salience adds to the overall interpretation of the scene by influencing which character is perceived as most important, but does not directly affect the accessibility of individual entities. This is

compatible with the view that the order of mention of characters in a scene is affected by a global apprehension of the event (cf. Griffin & Bock, 2000).

Secondly, the lack of an effect of linguistic salience on referent choice suggests that in Experiment 2 the linguistic context did not imply a clear protagonist, because it did not explicitly introduce one of the characters as such (as opposed to *Er was eens…* 'Once upon a time there was…' in Experiment 1). Thus, in Experiment 2, linguistic salience was solely defined in terms of subjecthood, which did not, as in Experiment 1, affect the choice of referent. This would also support the view that referent choice is affected more by global aspects of the narrative than by local saliency. Another explanation for the absence of a linguistic salience effect could be that it is due to the large preference of participants to mention the agent character first. This might have overridden any effects of the linguistic context. However, it is not clear why this preference did not (completely) override the effect of visual salience in Experiment 2, or the effect of linguistic salience in Experiment 1, in which there was also a large preference to start the sentence with the agent character.

### 2.4. General Discussion

Two story completion experiments showed that an entity's visual salience, as measured by its position in the foreground or in the background of the scene, influences the choice of referent, but not the choice of referring expression. Participants were more likely to start their utterance with the agent character when this character was visually salient, but they were not more likely to refer to this character with a pronoun. In addition, visual salience did not interact with linguistic salience, i.e., whether the referent was the subject of the previous sentence. In contrast to visual salience, linguistic salience affected the choice of referring expression, but influenced referent choice only in Experiment 1, where participants were more likely to take the subject of the first context sentence as the protagonist.

The results of the present study follow up on earlier findings in the following ways. Firstly, the finding that the choice of referent is influenced by the perceptual properties of the characters supports findings from scene description tasks that showed that visual and attentional cues influence order of mention (Gleitman et al., 2007; Osgood, 1971; Osgood & Bock, 1977; Tomlin, 1997). Whereas such tasks typically elicited isolated sentences, we have shown that these effects also apply in narrative

discourse contexts, in which the visual entities being mentioned are also available from the preceding linguistic context. This means that an entity's linguistic salience does not override effects of perceptual salience. In fact, only when one of the characters was linguistically marked as a protagonist (using the phrase *Er was eens…* 'Once upon a time there was…', Experiment 1), the linguistic context had an influence on the choice of referent. This mirrors the effect of linguistic salience in Prat-Sala and Branigan (2000), who also found that the entity that had been introduced in an existential construction (*There was…*) was more likely to be referred to first. In line with Prat-Sala and Branigan's findings, our results support the view that it is a combination of linguistic and non-linguistic factors that determines order of mention.

Secondly, our results support earlier findings (e.g., Arnold, 1998; Gordon et al., 1993; Stevenson et al., 1994) that a referent's salience in the linguistic context has an impact on the choice of referring expression: The likelihood of choosing a pronoun over a full noun phrase is higher when the referent is the subject of the directly preceding sentence than when it is not mentioned in the previous sentence. More importantly, our results show, in contrast to Fukumura et al. (2010), that the choice of referring expression is not affected by the salience of the referent in the visual context, even when the referent's salience from the linguistic context is unclear. Our manipulation of visual salience was different from that by Fukumura and colleagues. While they varied the number of possible referents in the scene, our scenes always involved the same number of characters, but varied in the position of the characters relative to the observer (or, more precisely, the camera). In visual perception, foregrounded objects have been found to be attended to more than objects in the background (Mazza et al., 2005). This might be a more sound manipulation of salience, since the one employed by Fukumura et al. also changes the number of items that have to be kept in memory, affecting cognitive load, which may in turn affect the choice of referring expression (cf. Arnold & Griffin, 2007).

Our finding that visual salience affects the choice of referent but not the choice of referring expression supports the hypothesis that people use different types of information in choosing a subject referent and choosing a referring expression. The fact that visually foregrounded characters are more likely to be referred to as subjects indicates that the spatial location of the characters in the pictures is a salient feature that influences language production. However, this information is apparently not used in choosing a referring expression. Thus, our results suggest that the choice of referring expression should be dissociated from the choice of referent. While the

choice of referring expression is likely to be dependent on the accessibility of mental representations (Ariel, 1990; Gundel et al., 1993), this does not automatically imply that the same information is also used to pick a referent for first mention. Rather, referent choice may be affected by protagonist or figure-ground assignment, which may be separate from the accessibility of the mental representations associated with individual entities (cf. Gleitman et al., 2007). In our experiments, people may have been more inclined to take the perspective of the visually foregrounded character. Hence, they may have interpreted it as the main character or figure in the discourse, which made it more likely that they referred to this character first. This account is consistent with the view that speakers primarily plan their utterances based on the global structure of an event, e.g., concerning the relations between the entities in the scene, and less by taking into account the salience of individual elements (Bock et al., 2004; Griffin & Bock, 2000).

Similarly, in the linguistic context, the character introduced as the protagonist is more likely to be mentioned as the subject in the continuation, whereas the choice of referring expression is more dependent on the most recent subject (Gordon et al., 1993; Stevenson, 2002). While it has been found that protagonisthood may also affect the choice of referring expression (e.g., Anderson et al., 1983; Karmiloff-Smith, 1981), our results suggest that local discourse salience is a stronger factor here (cf. Van Vliet, 2008). This might be an indication that referent choice is primarily determined by global conceptual aspects of the narrative, such as which is the main character in the discourse, while the choice of referring expression seems to depend more on the model the speaker has of the preceding discourse, such as the referent's accessibility status in the immediate linguistic context.

The discrepancy that we find between the effects of linguistic and visual salience might originate in the way salience is defined. One could argue that salience in the visual context and salience in the linguistic context are in fact two different notions of salience. While linguistic salience may be seen as a property of an entity's representation in the discourse, visual salience is a property of the stimulus itself. Linguistic salience is often associated with topicality and predictability (e.g., Givón, 1983; Arnold, 2008, 2010): Salient referents are those referents that are likely to be mentioned again in the discourse, given their discourse status. Since predictable referents are more accessible, they are also more likely to be referred to with attenuated expressions. Visually salient entities, on the other hand, may attract attention, but it is unclear whether this necessarily increases the accessibility of their

representations (cf. Arnold, 2010). It might even be conceivable that visual salience, at least in the way it was manipulated in the present study, is more associated with focus and unpredictability: It could be the case that visually foregrounding an entity marks it as important and having high news value. Similarly, visually backgrounding an entity might correspond to the linguistic notion of background, which contains given, already established information (e.g., De Swart & De Hoop, 2000). We leave the discussion of the exact informational status of visual salience to further research. In addition, it remains to be seen how manipulating visual salience using other perceptual properties, such as a referent's inherent size, its color, dynamics, or affective properties, might affect reference.

Another possible explanation for the fact that we did not find that visually salient entities are referred to with more attenuated expressions is that visual salience might be too subtle to affect the choice between nominal and pronominal expressions. It could be the case that this division is too rough. Effects of visual salience might be only visible on a more fine-grained scale. To test this possibility, we investigated the use of reduced pronouns in Experiment 1 (Experiment 2 contained too few reduced pronouns to perform statistical analyses). Our experiments were conducted in Dutch, which has a distinction between full (*hij* 'he', *zij* 'she') and reduced (*ie*/*die* 'he', *ze* 'she') pronouns. Because the syntactic distribution of masculine reduced pronouns is more restricted than that of the feminine forms (they cannot occur sentence-initially), we only investigated the feminine reduced pronouns, which were compared to the feminine full pronouns. We found no significant effects of either linguistic or visual salience on the use of reduced pronouns, suggesting that salience does not play a large role at this level.

A further issue concerns the linguistically undetermined conditions in Experiment 2. These were created to make sure that effects of linguistic salience would not dominate those of visual salience. However, there was still a strong preference to use a full NP when referring to the agent character. This preference could have obscured effects of visual salience. It is possible that the preference to use a full NP is due to competition between the two characters. In the linguistically undetermined conditions, both characters were made subjects in the preceding context. It was predicted that they would be equally salient in the linguistic context, making the relative salience of the two characters unclear. Because the two characters are salient in the mind of the speaker, a large amount of attention may be allocated to both of them. However, if attention is a limited resource, the amount of attention captured by each entity should

be lower than when all attention can be allocated to one salient character (Arnold & Griffin, 2007). Hence, the accessibility of each character might actually be lower, which results in more elaborate referring expressions, such as full NPs.

It could also be the case that because the context sentences in our experiments were presented auditorily and the participants had to respond orally within a fixed time frame, participants had a hard time keeping track of the characters mentioned in the linguistic context. As a result, they might have taken the last mentioned character as the most salient, i.e., of which activation was not yet decaying (e.g., Almor, 1999). To test this, we conducted another experiment with the same materials as Experiment 2, but with written rather than spoken context sentences. In addition, participants had to write their response down while they could still view the context sentences and could take all the time they liked. The results of this experiment were very similar to that of Experiment 2, with only slightly more pronouns in the linguistically undetermined condition. This indicates that the extensive use of full NPs in this condition is probably not due to a fast decay of memory activations.

Nevertheless, both in Experiment 2 and in the written experiment, the proportion of pronouns used to refer to the agent character in the linguistically undetermined condition was significantly higher than that in the linguistically non-salient condition, in which the agent character had only been mentioned in a PP adjunct, and significantly lower than that in the linguistically salient condition, in which the agent character had been the only subject in the preceding two sentences. This indicates that the salience of the agent character in the linguistically undetermined condition was indeed less clear than in the other conditions.

## 2.5. Conclusion

In summary, the present study has shown that a character's visual salience in a narrative discourse guides speakers in choosing which character they refer to first, as the subject of their utterance, but not in what kind of referring expression they use. These results suggest that visual salience does not affect the accessibility of individual entities. Rather, it influences the global interpretation of the situation. Our results are compatible with the view that utterance planning is not directly influenced by low-level perceptual properties, but that at least some higher-level conceptual information is required for the formulation of a message. Our results also suggest that the choice of

referring expression is more dependent on the status of the speaker's current discourse model, and hence works on a more local level.

# Chapter 3
# Lexical and perceptual animacy

**Abstract**

Several studies suggest that referential choices are influenced by animacy. On the one hand, animate referents are more likely to be mentioned as subjects than inanimate referents. On the other hand, animate referents are more frequently pronominalized than inanimate referents. These effects have been analyzed as effects of conceptual accessibility. In this chapter, we raise the question whether these effects are driven only by lexical concepts, such that referents described by animate lexical items (e.g., 'toddler') are more accessible than referents described by inanimate lexical items (e.g., 'shoe'), or whether they can also be influenced by context-derived conceptualizations, such that referents that are perceived as animate in a particular context are more accessible than referents that are not. In two animation-retelling experiments, conducted in Dutch, we investigated the influence of lexical and perceptual animacy on the choice of referent and the choice of referring expression. If the effects of animacy are context-dependent, entities that are perceived as animate should yield more subject references and more pronouns than entities that are perceived as inanimate, irrespective of their lexical animacy. If the effects are tied to lexical concepts, entities described with animate lexical items should be mentioned as the subject and pronominalized more frequently than entities described with inanimate lexical items, irrespective of their perceptual animacy. The results show that while only lexical animacy appears to affect the choice of subject referent, perceptual animacy may overrule lexical animacy in the choice of referring expression. These findings suggest that referential choices can be influenced by conceptualizations based on the perceptual context.

**This chapter is based on:**

Vogels, J., Krahmer, E. J., & Maes, A. A. (2013). When a stone tries to climb up a slope: the interplay between lexical and perceptual animacy in referential choices. *Frontiers in Psychology*, *4*:154.

### 3.1. Introduction

Throughout the languages of the world, the influence of animacy turns up in numerous linguistic choices. For example, animate entities are more likely to be chosen as the subject or the topic of a sentence than inanimate entities (e.g., Dahl & Fraurud, 1996; Givón, 1983), and they also typically occur earlier in the sentence (e.g., Branigan & Feleki, 1999). The tendency to place animate entities early in the sentence also leads to animacy effects in the choice between alternating grammatical structures. For example, passive sentences are more frequent when the patient role is taken up by an animate entity (McDonald, Bock, & Kelly, 1993; Van Nice & Dietrich, 2003a). This is illustrated by the preference in English for the sentence in (1a) over the one in (1b) (Prat-Sala & Branigan, 2000).

(1) a. *The woman was run over by the train.*
b. *The train ran over the woman.*

In addition, there is evidence that animacy affects the choice of referring expressions: Animate entities have been found to be more often referred to with pronouns than inanimate entities (Dahl & Fraurud, 1996; Fukumura & Van Gompel, 2011; Yamamoto, 1999). For example, Fukumura and Van Gompel (2011) found in a story completion experiment that speakers were more likely to pronominalize the animate entity (*the hikers*) than the inanimate entity (*the canoes*) in (2a). The same held when grammatical roles were reversed, such as in (2b), suggesting that the effect of animacy on pronominalization is independent of grammatical function. They also found that animacy affected the choice of referent: Participants were more likely to refer to the animate NP than to the inanimate NP in their continuations.

(2) a. *The hikers carried the canoes downstream. Sometimes...*
b. *The canoes carried the hikers downstream. Sometimes...*

Thus, animacy appears to influence referential choices: On the one hand, it affects which referent is chosen as the subject of the sentence or as the first-mentioned entity. On the other hand, it affects the type of referring expression that is used to refer to an entity, e.g., a pronoun (‘she’) or a full noun phrase (‘the girl’). These effects are generally explained as effects of conceptual accessibility (Bock & Warren, 1985):

Mental representations of animate entities are more easily retrieved from memory than representations of inanimate entities. Therefore, they are available early for linguistic processing (e.g., Prat-Sala & Branigan, 2000), and need less linguistic encoding (Fukumura & Van Gompel, 2011). However, it is less clear what the source of these effects is. It could be the case that they arise from the accessibility of lexical concepts (e.g., Branigan, Pickering, & Tanaka, 2008), such that the representations associated with animate nouns in the mental lexicon (e.g., 'toddler') are more accessible than those associated with inanimate nouns (e.g., 'shoe'). Alternatively, the effects of animacy could be driven by the accessibility of non-linguistic conceptual representations, which may be influenced by the (perceptual) context (e.g., Arnold, 2010; Bock & Warren, 1985). For example, entities that move in a (seemingly) meaningful way may be conceptualized as more animate, and therefore be more accessible. In this chapter, we investigate the interplay between a referent's animacy based on the associated lexical concept (lexical animacy) and the degree to which it is conceptualized as animate or inanimate based on motion cues (perceptual animacy) in referential choices. We investigate effects on both referent choice (which entity is referred to as the subject), and choice of referring expression (whether the entity is referred to with a pronoun or a full noun phrase) in Dutch.

### 3.2. Theoretical background

Although strictly speaking the term 'animacy' refers to the degree to which something is alive, animacy is not a property of entities in the world. Rather, it is a property of people's *cognitive representations* of entities, which result from the way people mentally classify entities in the world as 'animate' or 'inanimate' (a cognitive ontology, Fraurud, 1996). Therefore, this classification differs from a strictly biological sense of 'livingness'. For example, in the animacy hierarchy given in (3) (e.g., Comrie, 1989), humans are treated as more animate than animals, although they are not more 'alive' in a biological sense.

(3) *Animacy hierarchy*
Human > Animate > Inanimate

In addition, the way animacy has been found to affect linguistic structure shows that animacy can be a more gradient factor than suggested by the hierarchy in (3). For example, entities such as machines and vehicles, or collectives such as companies and organizations, are treated linguistically as more animate than objects like books and tables (e.g., Comrie, 1989; Dabrowska, 1998; Rosenbach, 2008). Thus, what counts as more animate or inanimate is not so much dependent on properties intrinsic to entities, but on how we *conceptualize* these entities.

In early transformational grammar (e.g., Katz, 1972), animacy was formalized as a semantic feature tied to an entity's lexical item. A feature that did not match the selection restrictions evoked by the predicate would result in an anomaly. Hence, the sentence in (4) would be anomalous, since 'chase' takes an animate subject, while 'tree' does not have the feature 'animate'.

(4) **The tree chased the fly.*

Although the anomaly might be resolved in certain contexts, the structure in itself remains ungrammatical under this account. Thus, animacy is regarded here as closely tied to the lexicon. More recently, many (psycho)linguistic studies on animacy also treat the conceptualization of entities only implicitly, presupposing an animate representation for an animate lexical item (e.g., 'toddler'), and an inanimate representation for an inanimate lexical item (e.g., 'shoe'). For example, Prat-Sala and Branigan (2000) assume that an entity's animacy contributes to its inherent salience, which is constant across contexts. In addition, Branigan et al. (2008), while considering animacy as one of the factors affecting a referent's conceptual accessibility, assume that this refers to the accessibility of *lexical concepts*, i.e., concepts that are closely connected to a lexical item.

However, it is clear that people do not always assign the same degree of animacy to the same (lexical) concepts. Like discourse salience (e.g., topichood), this is something that can vary with context. Notably, in some contexts people can conceptualize usually inanimate entities as animate. In cartoons or fairy tales, for example, inanimate entities or animals are often anthropomorphized. This also happens in real-world contexts, as when someone says 'The tree wants to catch me' for a tree with branches sticking out like arms. The reverse, treating animate entities as inanimate, is

theoretically also possible, although this may be less likely.[1] In addition, in figurative language use such as personification, metaphor, and metonymy, entities are often referred to in a way that does not match their actual animacy, as in 'His ideas will live on forever' (Lakoff & Johnson, 1980), or 'The ham sandwich is sitting at table 20' (Nunberg, 1979), to refer to a customer in a restaurant.

Evidence that there is variation in the way entities are conceptualized as animate or inanimate comes from different areas of research. From linguistic typology we know that languages differ in which entities are treated as animate or inanimate in the grammar. A well-known example is that of the Algonquian language Fox, in which the word for 'strawberry' is grammatically inanimate, while the word for 'raspberry' is animate (Anderson, 1997). Hence, the former cannot occur as the beneficiary role in ditransitive constructions, while the latter can. Similarly, in Persian, the word for 'tree' is lexically classified as animate, by which it takes the animate plural suffix, while the word for 'flower' takes an inanimate suffix (Wiese, 2003). In addition, in many European languages inanimate nouns have masculine or feminine gender, which may affect how they are conceptualized (e.g., Boroditsky, Schmidt, & Phillips, 2003; Dahl, 2000).

There is also evidence that conceptualizations of animacy may differ across contexts within the same language. In an ERP-study by Nieuwland and Van Berkum (2006), utterances that would normally violate animacy requirements, such as 'The peanut was in love', were found to be easy to process when they were embedded in a fairy tale or cartoon-like context, in which inanimate objects were consistently made the subject of predicates that require an animate subject (e.g., 'dance'). Within such contexts, a normally well-formed utterance such as 'The peanut was salted' became more difficult to process. This suggests that people can easily accommodate to contexts in which normally inanimate objects are presented as animate, and that these objects are treated as fully animate from that point on. This makes it clear that the classification of concepts according to their lexical-semantic animacy should be distinguished from contextually inferred animacy (cf. Rosenbach, 2008; Yamamoto, 1999).

[1] One reason why such cases are hard to find may be that there are more linguistic constructions that require animate arguments than there are constructions that only take inanimate arguments. An example might be the use of animate nouns in constructions that normally only allow mass nouns, as in *That's a lot of dog you've got there* (Croft, 1994).

The conceptualization of entities as more or less animate in reference may be related to the anthropocentric nature of language, i.e., people talk about things from their own, human, perspective. An entity may thus be more animate the more it resembles humans. The reason for this may be that people have more empathy towards such entities (e.g., Kuno & Kaburaki, 1977), or find them otherwise more important to talk about (Givón, 1983). Sridhar (1988), for example, found that when people described an interaction between a ball and a doll (two inanimate objects), the doll was more likely to be mentioned first in the sentence. It has been suggested that the relevant property in conceptualizing referents is degree of individuation, e.g., whether they appear to be autonomous beings, have the ability to act upon their environment, or have goals, intentions and mental states (Dahl, 2008; Fraurud, 1996).

An important factor in classifying an entity as an individual is how it is perceived. For example, individuals typically exist in their own right (they are not physically part of another entity), they move without the intervention of an external force, and they act in meaningful ways. Hence, a clear perceptual cue for animacy or individuation is motion. According to the perception literature, movements of simple geometric objects can indeed induce a strong and immediate percept of animacy (e.g., Scholl & Tremoulet, 2000). In an early study by Heider and Simmel (1944), participants were found to readily assign emotions and intentions to geometric objects when they moved in non-random ways. More recent work shows that even very subtle movement cues can still create a perception of animacy. For example, a sudden change in speed or direction already leads to animate percepts (Tremoulet & Feldman, 2000). In addition, when the movements of two objects are correlated or one moving object pauses near another object, this creates a suggestion of animacy (Santos, David, Bente, & Vogeley, 2008; Schultz, Friston, O'Doherty, Wolpert, & Frith, 2005).

Applied to normally inanimate entities, such perceptual motion cues may cause them to be conceived of as more animate and more individuated, which may make them more conceptually accessible. According to the theory of conceptual accessibility (Bock & Warren, 1985), the activation of mental representations of referents in memory is fed by both perception and conceptual knowledge. Indeed, Bock, Loebell, and Morey (1992) and McDonald et al. (1993) found that animacy effects on word order were enhanced when participants created a mental image of the entities. Similarly, in the Nieuwland and Van Berkum (2006) study most participants reported to have visualized the story and to have seen the inanimate objects as cartoon-like characters with human characteristics such as a face, arms and legs. This suggests that

context-dependent perceptual information, such as motion, can contribute to a referent's conceptual accessibility, on top of conceptual information from long-term memory.

Prat-Sala and Branigan (2000) distinguish two types of accessibility: inherent accessibility, which concerns properties of a referent that remain stable across contexts, and derived accessibility, which concerns the salience of a referent in the linguistic or non-linguistic context. Factors influencing inherent accessibility typically include lexical animacy (e.g., Prat-Sala & Branigan, 2000) and concreteness (Maes, 1997), while derived accessibility is typically affected by factors such as givenness (e.g., Ferreira & Yoshita, 2003) and thematic role (Van Nice & Dietrich, 2003b). Since perceptual motion cues for animacy may change across contexts, and are not intrinsic to the entity itself, we may add perceptual animacy as a factor contributing to derived accessibility.

In Prat-Sala and Branigan's (2000) view, a referent's overall conceptual accessibility is a combination of its inherent and its derived accessibility. They also argued, on the basis of a picture description experiment, that in strong enough contexts, derived accessibility might override effects of inherent accessibility. Hence, a lexically animate entity is more likely to be mentioned in subject position than an inanimate entity, unless the discourse makes the inanimate entity salient enough to overcome the difference in inherent accessibility. Additional evidence that derived accessibility may override inherent accessibility has been found, e.g., by Christianson and Ferreira (2005) and Van Nice and Dietrich (2003b). This is also consistent with the findings of Nieuwland and Van Berkum (2006), who found that the lexical meaning of 'peanut' was overruled by the pragmatic inference of the referent's animacy due to the discourse context. On the other hand, other studies have not found evidence for the dominance of one type of accessibility over the other (e.g., Fukumura & Van Gompel, 2011; Van Nice & Dietrich, 2003b).

The question we ask in this chapter is how lexical-semantic animacy and contextually driven animacy interact in determining referential choices in language production. To address this question, the present study investigates the interplay between the animacy associated with lexical concepts (henceforth *lexical animacy*), and the perceived animacy based on the referent's movements (henceforth *perceptual animacy*) in Dutch spoken language production. That is, we investigate whether lexically inanimate referents that are conceptualized as animate and lexically animate referents that are conceptualized as inanimate are different from the congruent cases

with respect to referential choices. We examine both the choice of referent (which entity is referred to as the subject) and the choice of referring expression (use of pronouns and full noun phrases).

One possible hypothesis is that lexical and perceptual cues for animacy both affect the conceptual accessibility of a referent, but that perceptual animacy overrules lexical animacy, in line with Nieuwland and Van Berkum (2006). Hence, animate moving objects should be more likely to be referred to as subjects and with pronouns than inanimate moving objects, irrespective of lexical animacy. Only when the entity is perceptually inanimate, lexical animacy is expected to have an effect, since conceptualizing animate entities as inanimate may be less straightforward.

Alternatively, lexical and perceptual animacy may affect accessibility independently. In this case, a referent's accessibility is predicted to be highest when the entity is both mentioned using an animate lexical description and perceived as animate. It should be lowest when the referent is both lexically and perceptually inanimate. In the incongruent cases, i.e., lexically animate but perceptually inanimate or vice versa, accessibility is predicted to be intermediate. Assuming that both the rate of pronominalization and the likelihood of being mentioned as the subject increase proportionally with an increase in accessibility, pronouns and subject references are predicted to be most frequent in cases where all cues point to a high degree of animacy, to be least frequent in cases where all cues point to a low degree of animacy, and somewhere in between for the incongruent cases. If conceptualizing lexically inanimate objects as animate is easier than conceptualizing lexically animate objects as inanimate, the effect of perceptual animacy should be at least present in the lexically inanimate condition.

We conducted two experiments, in which participants watched animations of simple geometric objects, such as circles and triangles, and retold them afterwards. We used retelling from memory rather than speaking when the animations were still in view because we believed this would be the more natural communicative situation (cf. Christianson & Ferreira, 2005). The perceptual animacy of the objects in the animations was manipulated by using movement cues to create animate and inanimate conceptualizations. Manipulating motion allowed us to make use of the exact same objects in the animate and in the inanimate conceptualizations. In this way, the appearance of the referent was kept constant across all conditions, such that it could not influence the referent's perceptual or lexical animacy. In addition, we separated lexical animacy from perceptual animacy by giving lexical labels to the

objects. In Experiment 1, the lexical labels were animate and inanimate nouns that either matched or did not match in animacy with the movements. In Experiment 2, we replaced the lexical labels with nonsense words that could be interpreted as either referring to animate or referring to inanimate entities, to exclude a possible influence of lexical animacy on perceptual animacy.

## 3.3. Experiment 1

### *3.3.1. Methods*

#### 3.3.1.1. Participants

Sixty-four students from Tilburg University participated in this experiment for course credit. All were native speakers of Dutch. All participants gave their consent to the use of their data.

#### 3.3.1.2. Materials

We created 16 different animations, using the motion paths from the custom animation function in Microsoft PowerPoint. Each animation featured three geometric objects, of which one was the target figure and the two others were competitors. The objects were selected from the following built-in shapes in PowerPoint: cross, oval, rectangle, isosceles triangle, up arrow, and diamond. All had the same dimensions. The two competitor objects both had the same shape, which was always different from the target figure's shape. The figures appeared in one of four colors: white, light green, light blue or light purple. Within one animation, colors of the target and the competitors were always the same. Shapes and colors were assigned randomly to the animations, except for animations involving rolling or bouncing movements, in which the target object was always a circle.

Eight animations contained animate motion of the target figure, and the other eight contained inanimate motion of the target figure (to be explained below). The animate and inanimate animations were paired, such that for each animate animation there was another animation featuring the same objects but in which the target figure moved in an inanimate way. An example of an animate stimulus item is given in Figure 3.1. The animations were presented on a black background, but some animations included the suggestion of a landscape, presented by a white continuous

line (as in Figure 3.1). This was done to aid the interpretation of some movements (e.g., 'climbing up a slope' instead of 'taking off magically into nothingness').

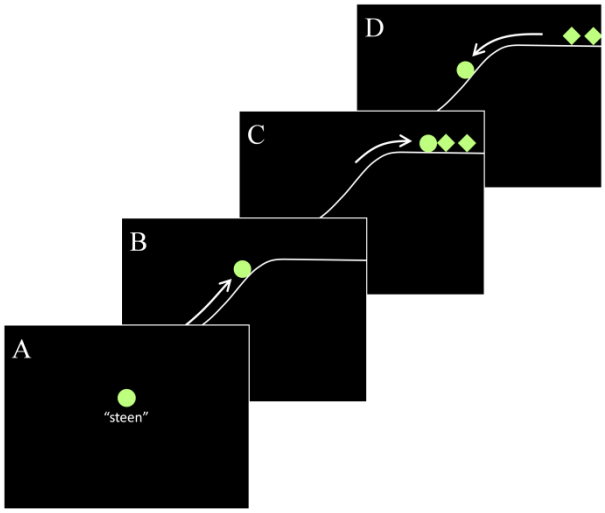


**Figure 3.1.** Example of an incongruent stimulus item in Experiment 1, with the target figure moving in an animate manner, but having the inanimate lexical label *steen* 'stone'. The four frames (A-D) are stills taken from a continuous animation. Letters indicate order; arrows indicate movement. Both were not shown in the experiment.

Each target figure was given a linguistic label (in Dutch), either animate (e.g., *padvinder* 'boy scout') or inanimate (e.g., *steen* 'stone'). The animate and inanimate labels were matched (across items) for frequency and number of characters. The complete list of lexical labels can be found in Table A1 in Appendix A. The competitors had no labels. In each trial, the target figure was presented just before the start of the animation along with its label (frame (A) in Figure 3.1). To make repeated references possible, each animation consisted of three 'episodes'. First, the target figure performed an intransitive action (B). Here, the target figure moved either in an animate way or in an inanimate way. Animate movements were suggested by simulating self-propelled actions (e.g., climbing up a slope), using cues such as changes in speed or direction (Tremoulet & Feldman, 2000). The animate movements used were: moving back and forth horizontally across the screen at varying speeds (2 animations); jumping up and down irregularly; hopping back and forth across the screen with varying intervals (2 animations); moving up a slope with a pause just before the top; moving diagonally up across the screen in small steps; making irregular arc-shaped movements across the screen. Inanimate movements were

suggested by creating the impression that they were caused by an external (invisible) force such as gravity (e.g., rolling down a slope; Gelman, Durgin, & Kaufman, 1995). Since we wanted to keep implicit what set the object in motion, inanimate movements necessarily started off-screen. The inanimate movements used were: moving across the screen at a constant speed (i.e., as if sliding on ice; 2 animations); bouncing vertically several times with loss of energy; bouncing once (2 animations); rolling down a slope; whirling from top to bottom; moving down and up the slopes of a valley with loss of energy.

Second, the competitor figures entered the screen, and the target figure performed a transitive action (interaction with competitor figures (C)). Animate movements included: colliding with the competitors from rest; jumping on and off the competitors (2 animations); pushing the competitors away (2 animations); quickly jumping up and down in front of the competitors (as if startled); bumping into the competitors; briefly touching the competitors. Inanimate movements included: colliding with the competitors and bouncing back (2 animations); land on top of the competitors (3 animations); colliding with the competitors and pushing them away (2 animations); bouncing over the competitors. The competitors were included to allow alternating syntactic structures (e.g., 'the stone hits the two hikers' vs. 'the two hikers are hit by the stone'), as well as to encourage reference switches, which should lead to variation in referring expression use (both pronouns and full noun phrases). They appeared in dyads, such that pronominal references to the target or the competitors were likely to be unambiguous (singular vs. plural). The only movement that the competitors made was sliding into the screen (either from the left or the right). Since this movement was not particularly animate or inanimate, the perceptual animacy of the competitors remained ambiguous. The target figure always appeared before the competitors, to make it a likely candidate for the discourse topic. To control for agency, the target figure was always the agent in the transitive action, both in the animate and in the inanimate conditions.

Finally, the target figure performed another intransitive action (D). Animate movements included: (quickly) moving off the screen from rest (4 animations); hopping off the screen from rest; rolling down the slope, off the screen; quickly hopping or stepping down diagonally, off the screen (2 animations). Inanimate movements included: moving off the screen; being bounced off the screen; landing on the ground while turning on its axis (3 animations); bouncing back and coming to a rest (3 animations).

The animations (without the lexical labels) were pretested for perceived animacy of the target referent in a perception study. Eight participants were asked to rate the target referent in each of the animations for animacy on a seven-point Likert scale, with 1 being 'clearly lifeless' and 7 being 'clearly alive' (cf. Tremoulet and Feldman, 2000). Ratings were given for each animation as a whole, not for each movement separately. The results confirmed that animations intended to be animate were scored significantly higher ($M_{anim}$ = 5.13; $M_{inan}$ = 2.53; One-way ANOVA: $F(1,126) = 114.48$, $p < .001$, $MSE = 1.88$).

3.3.1.3. Procedure

The participants were seated at a table, facing the experiment leader, who sat in a chair facing the participant. The experiment was presented on a laptop, which was on the table at an angle with the participant. The participants' task was to retell the animations to the experiment leader. Before each animation, the target figure was presented in the middle of the screen, accompanied by its lexical label. The target figure and the label were then replaced by a crosshair, after which the animation started. Participants watched each animation twice, so that they could accurately retell them from memory. They were not allowed to start talking when the animation was still running, because this might have caused them to skip over crucial information.

The participants were instructed to use the label presented in the beginning when mentioning the target figure. (Of course, participants were allowed to pronominalize referents. Although it was not instructed explicitly, all participants did this.) The competitors could be referred to in any way they wanted. To ensure lively retellings, participants were further instructed to retell the animations 'in a fanciful manner, as if telling it to a child'. The experiment started with three practice trials, after which any remaining questions could be asked. The experiment leader gave only minimal feedback during the experiment (e.g., nodding or saying 'Okay' after each trial; in a few cases the participant received some encouragement to start talking). There were no further interactions between participant and experiment leader while the experiment was running. It took about 25 minutes to complete the experiment.

3.3.1.4. Design

Crossing the factors lexical animacy and perceptual animacy resulted in a 2 (lexically animate, lexically inanimate) x 2 (perceptually animate, perceptually inanimate) within-participants and within-items design. Participants were randomly assigned to

one of four lists, which were created such that from a given item each condition occurred on a different list. The items were presented in a random order. The same order was used across all lists.

3.3.1.5. Data coding and statistical analyses

The data from one participant were discarded, because this person retold the animations while watching them instead of afterwards. The data from the remaining participants were coded by the first author. Uncertainties were resolved through discussion with the other authors. First, all stories were divided into fragments containing descriptions of the three episodes (initial intransitive action, transitive action, and final intransitive action). We focused on the descriptions of the transitive action, since these were the fragments that were expected to show most variation in choice of referent for the subject position and choice of referring expression.

Next, the fragments were coded for whether the target figure was made the subject of the critical clause (referent choice), and whether the target figure was referred to using attenuated expressions (choice of referring expression). We coded all grammatical subjects of both main and subordinate clauses as 'subject', and everything else was coded as 'object'. We defined attenuated expressions as all referring expressions that were not full noun phrases. These included full pronouns (e.g., *zij* 'she'), reduced pronouns (e.g., *ze* 'she'), demonstrative pronouns (e.g., *die* 'that one') and zero anaphora (e.g., *...en Ø springt over twee huizen* '...and Ø jumps over two houses'). Henceforth, we will use the term 'pronoun' to refer to all these types of referring expressions. If there was more than one clause describing the action, we only coded the first one. Trials in which the transitive action was not described were excluded from analysis. In addition, we excluded trials in which reference was made to a person or object that was not present in the animation, since this could have altered the interpretation of the referent's animacy, as well as its discourse accessibility. Finally, we excluded trials in which the target figure was referred to with an indefinite NP (generally used in contexts where a pronoun is not possible), referred to with a proper noun (typically used for animate referents), or not referred to at all. In all, 96 cases (19.0%) were removed. We controlled for discourse salience by coding whether the target figure was mentioned in the sentence directly preceding the clause under consideration, and if so, in what grammatical function.

We analyzed the data using logit mixed models (Jaeger, 2008). Lexical animacy and perceptual animacy were included as fixed factors; participants and items were

included as random factors. Fixed factors were centered to reduce collinearity. Two analyses were carried out: one on the log odds of a subject reference, and one on the log odds of a pronoun reference. Starting with the full random effect specification, we omitted random slopes that did not significantly affect model fit using model comparison (Jaeger, 2011). We present only the final models.

### *3.3.2. Results*

#### 3.3.2.1. Choice of referent

Figure 3.2 shows the proportion of references to the target figure as the subject of the clause describing the transitive action. Lexically animate target figures were mentioned as the subject in 91.0% of the cases, whereas lexically inanimate target figures were mentioned as the subject in 80.7% of the cases. The effect of lexical animacy on choice of referent was significant, $\beta = 0.93$, SE = 0.31, $p < .01$. We found no effect of perceptual animacy, $\beta = 0.29$, SE = 0.32, $p = .37$, and no interaction, $\beta = 0.28$, SE = 0.62, $p = .65$. Random slopes were not included, since they did not improve model fit.[2]

[2] It is conceivable that the effect of lexical animacy is due to relative animacy of the target and competitors rather than target animacy alone. Since participants were free to refer to the competitor objects in any way they wanted, we did not manipulate competitor animacy systematically. However, to check whether relative animacy could have affected our results, we coded whether participants used animate or inanimate lexical items to refer to the competitor objects. We omitted an additional 10 observations in which the animacy of the competitors was unclear. Adding competitor animacy as a factor to the model revealed a main effect of competitor animacy: The target was less likely to be mentioned as the subject when the competitors were lexically animate (72.7%; n = 150) than when they were lexically inanimate (94.4%; n = 248), $\beta = -2.28$, SE = 0.42, $p < .001$. The main effect of target lexical animacy remained significant, $\beta = 1.43$, SE = 0.46, $p < .01$, and target perceptual animacy remained non-significant, $\beta = 0.33$, SE = 0.48, $p = .50$. There were no interactions between target and competitor animacy, suggesting that target animacy affects subject mention independently from competitor animacy.

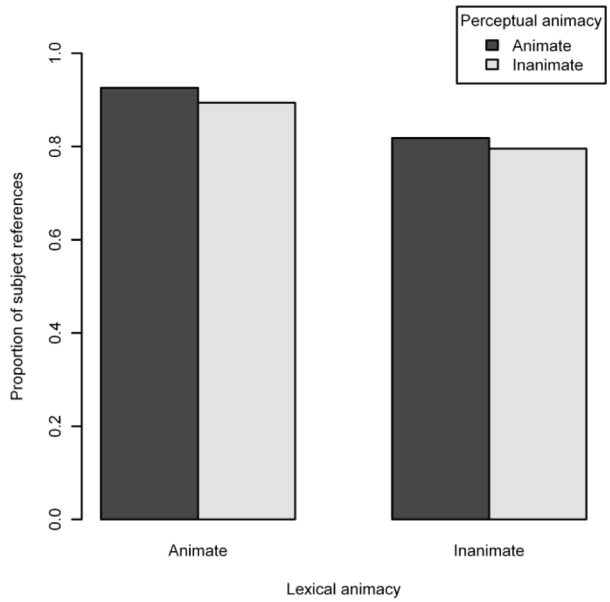


**Figure 3.2.** Proportion of subject references to the target figure in the description of the transitive action in Experiment 1, by its lexical and perceptual animacy.

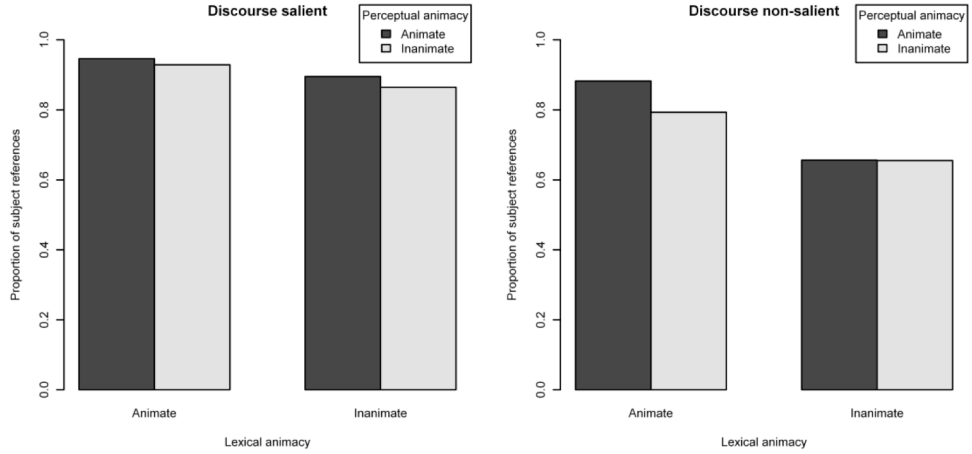


**Figure 3.3.** Proportion of subject references to the target figure in the description of the transitive action in Experiment 1, by its lexical and perceptual animacy, and split by the discourse salience of the referent: mentioned in the directly preceding sentence (discourse salient, left pane), or not mentioned in the directly preceding sentence (discourse non-salient, right pane).

To examine whether the effect of lexical animacy was confounded with the salience of the target figure in the discourse, we performed separate analyses for the cases in

which the referent was mentioned in the directly preceding sentence (n = 284), and for the cases in which the referent was not mentioned in the directly preceding sentence (n = 124).[3] The results, presented in Figure 3.3, show that the effect of lexical animacy only holds when the referent was not mentioned in the previous sentence (discourse non-salient), $\beta$ = 1.07, SE = 0.45, $p < .05$. The difference between lexically animate and inanimate referents that were discourse salient was not significant, $\beta$ = 0.74, SE = 0.45, $p = .10$. Again, perceptual animacy was non-significant in both data sets, $\beta$ = 0.42, SE = 0.47, $p = .37$ and $\beta$ = 0.30, SE = 0.46, $p = .52$, respectively; interactions: $\beta$ = 0.70, SE = 0.91, $p = .44$ and $\beta$ = -0.03, SE = 0.91, $p = .97$, respectively.

3.3.2.2. Choice of referring expression

Figure 3.4 shows the overall proportion of pronoun (i.e., non-full NP) references to the target figure in the description of the transitive action. This includes both subject and object pronouns. Lexically animate referents were referred to with pronouns in 86.0% of the cases, against 78.6% for lexically inanimate referents. The effect of lexical animacy was significant, $\beta$ = 0.58, SE = 0.28, $p < .05$. In addition, perceptually animate referents were referred to with pronouns in 85.5% of the cases, against 79.6% for perceptually inanimate referents. Although this effect was just slightly smaller than that of lexical animacy, it was only marginally significant, $\beta$ = 0.48, SE = 0.28, $p = .09$. We found no interaction between lexical and perceptual animacy, $\beta$ = -0.22, SE = 0.55, $p = .69$. Random slopes were not included, since they did not contribute to the model's fit.[4]

[3] Although the grammatical function of the referent in the previous sentence is an important factor in determining whether it will be the subject of the next sentence (e.g., Grosz et al., 1995), we did not distinguish between different grammatical functions here, since there were only a few cases in which the referent was mentioned as something else than the subject (3.7%). Repeating the analysis using a distinction between subject references and everything else yielded similar results.

[4] Again, we conducted an analysis in which we included competitor animacy in the model. The effect of competitor animacy was marginally significant, $\beta$ = -0.53, SE = 0.32, $p = .09$, with slightly fewer pronominalized target referents when the competitors were lexically animate (80.7%; n = 150) than lexically inanimate (83.5%; n = 248). The significant effect of target lexical animacy remained, $\beta$ = 0.71, SE = 0.29, $p < .05$, as did the marginally significant effect of target perceptual animacy, $\beta$ = 0.53, SE = 0.30, $p = .07$. There were no interactions between target and competitor animacy. This suggests that target lexical animacy affects pronoun use independently from competitor lexical animacy.

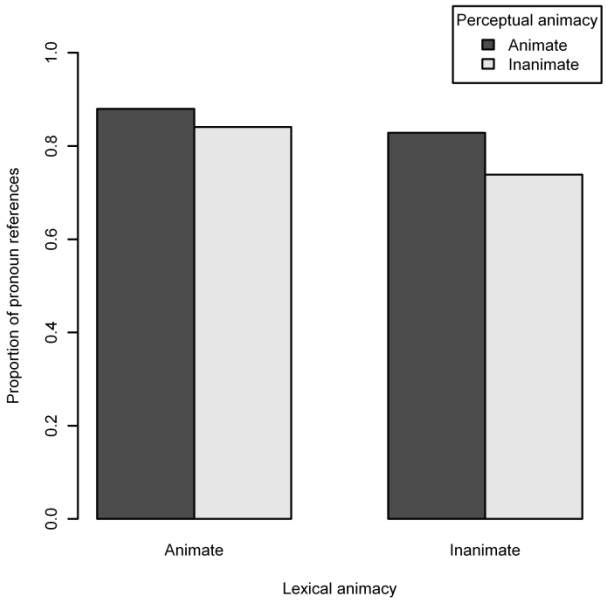


**Figure 3.4.** Proportion of pronoun references to the target in the description of the transitive action in Experiment 1, by its lexical and perceptual animacy.

To investigate whether these effects were confounded with the referent's discourse salience, we performed two separate analyses, one for the cases in which the referent was mentioned in the directly preceding sentence (n = 284), and one for the cases in which it was not mentioned (n = 124). The results, presented in Figure 3.5, show that when the referent was discourse salient (i.e., mentioned in the previous sentence), no effects of animacy were present (lexical animacy: $\beta$ = 0.90, SE = 0.60, $p$ = .14; perceptual animacy: $\beta$ = 0.31, SE = 0.62, $p$ = .62; interaction: $\beta$ = -0.05, SE = 1.20, $p$ = .97). However, when the referent was discourse non-salient (i.e., not mentioned in the previous sentence), perceptual animacy had a significant effect on the choice of referring expression, $\beta$ = 1.35, SE = 0.47, $p$ < .01: More pronouns were used when the referent was perceptually animate. The effect of lexical animacy was no longer significant, $\beta$ = 0.66, SE = 0.46, $p$ = .15, suggesting that this factor may indeed be partly confounded with discourse salience (i.e., what is lexically animate is also more likely to be the subject, cf. Figure 3.2). There was no interaction, $\beta$ = -0.58, SE = 0.91, $p$ = .52.

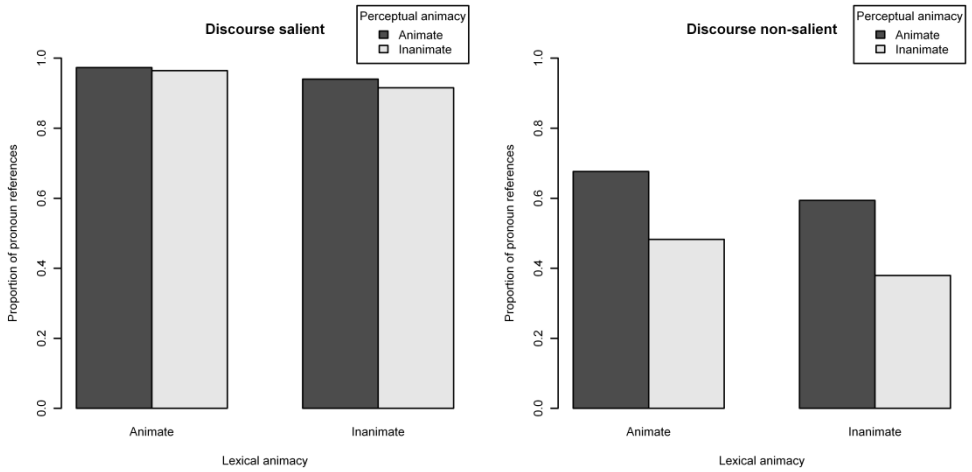


**Figure 3.5.** Proportion of pronoun references to the target figure in the description of the transitive action in Experiment 1, by its lexical and perceptual animacy, and split by the discourse salience of the referent: mentioned in the directly preceding sentence (discourse salient, left pane), or not mentioned in the directly preceding sentence (discourse non-salient, right pane).

We also investigated whether the grammatical function of the referent in the current sentence showed the same confound. To this end, we performed separate analyses on referring expressions in subject position (n = 352) and non-subject position (n = 56). The results, presented in Figure 3.6, showed a significant effect of perceptual animacy on the choice of referring expression in subject position, $\beta$ = 0.89, SE = 0.34, $p$ < .01: More pronouns were used when the referent was perceptually animate. There was no effect of lexical animacy, $\beta$ = 0.22, SE = 0.34, $p$ = .51, and no interaction, $\beta$ = -0.91, SE = 0.68, $p$ = .18. Although Figure 3.6 suggests an effect of perceptual animacy in the opposite direction in non-subject position, this was not significant, $\beta$ = -0.82, SE = 0.69, $p$ = .24. The same held for lexical animacy, $\beta$ = 0.99, SE = 0.67, $p$ = .14. There was no interaction, $\beta$ = 0.67, SE = 1.33, $p$ = .62. Again, these patterns suggest a confound of discourse salience/grammatical function with lexical animacy, but not with perceptual animacy.

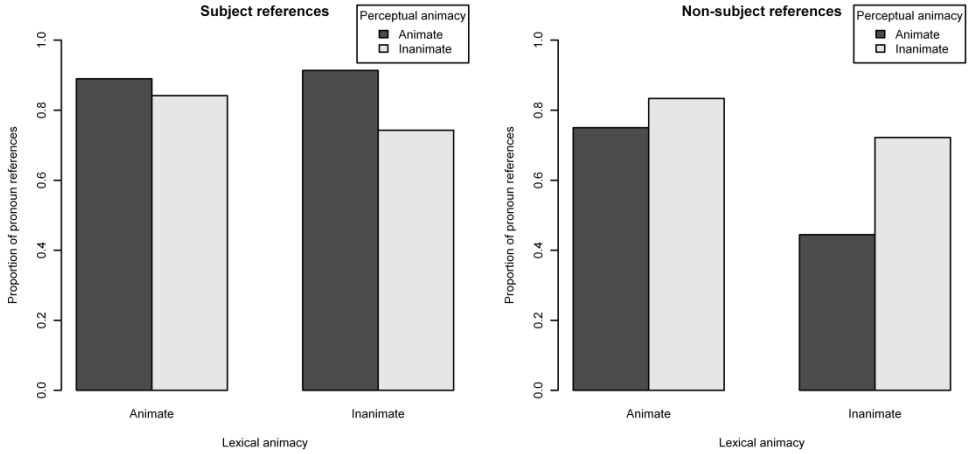


**Figure 3.6.** Proportion of pronoun references to the target figure in the description of the transitive action in Experiment 1, by its lexical and perceptual animacy, and split by the grammatical function of the referent: subject (left pane), or non-subject (right pane).

*3.3.3. Discussion*

The results of Experiment 1 showed differential effects of the two types of animacy manipulations. Firstly, in descriptions of the transitive event in the animations, which involved an interaction between the target figure and the two competitor figures, the target figure's lexical animacy, but not its perceptual animacy, influenced whether it was mentioned as the subject of the sentence or not. For example, participants were more likely to say *de padvinder duwt de pestkoppen weg* 'the boy scout pushes the bullies away' than they were to say *de steen duwt de pestkoppen weg* 'the stone pushes the bullies away' (in which case sentences such as *de pestkoppen worden weggeduwd door de steen* 'the bullies are pushed away by the stone' were more frequent).[5] However, whether it was an animate-like moving stone (e.g., one trying to climb up a slope) or not did not matter. In addition, the effects of animacy were only present when the referent was not mentioned in the previous sentence, suggesting that discourse salience is a stronger factor in determining which grammatical role is assigned to the referent.

Secondly, the choice of referring expression to refer to the target figure in the transitive action seemed to be affected by perceptual animacy: More pronouns were used when referents were moving in an animate way, independently of their lexical animacy. For example, participants were more likely to say *hij duwt ze weg* 'it pushes

[5] These examples, as well as those in the next paragraph, were constructed for illustrative purposes and do not reflect any participant's exact wordings.

them away' than they were to say *de steen duwt ze weg* 'the stone pushes them away', when the referent was an animate-moving stone. Again, this effect was confined to the cases in which the referent was discourse non-salient; when the referent was discourse salient, participants used pronouns almost exclusively, masking any effects of animacy. In addition, the effect of perceptual animacy was also confined to the cases in which the referent was the subject of the current sentence, suggesting that the environments in which perceptual animacy becomes relevant are cases of topic shift. Although pronominalization of lexically animate referents was also somewhat more frequent than pronominalization of lexically inanimate referents, this difference was not significant when controlling for discourse salience, suggesting that the effect of lexical animacy may be partly indirect, i.e., lexically animate referents are more likely to be subjects, and this in turn increases pronoun use. These findings show that the choice of referring expression can be influenced by factors induced by the non-linguistic context, such as perceptual animacy. Surprisingly, the effect of perceptual animacy was equally large in the lexically animate and lexically inanimate condition, suggesting that lexically animate referents can as easily be conceptualized as inanimate as lexically inanimate referents can be conceptualized as animate.

In Experiment 1, we tried to manipulate perceptual and lexical animacy independently. The results suggest that the two factors affect choice of referent and choice of referring expression differently. In the choice of referring expression, perceptual animacy seems to be a stronger cue for accessibility than lexical animacy, at least when discourse salience is low. Although lexical animacy does not seem to be completely overruled, this finding is in line with Nieuwland and Van Berkum's (2006) findings that in case of a conflict between the two factors, perceptual animacy gets prevalence. Thus, a stone that is trying to climb up a slope is assigned animacy because of its animate-like movements. In the choice of referent for the subject position, however, only lexical animacy seems to have an effect.

It is possible, however, that the lexical items influenced perception in the experiment, and that therefore the two factors were not independent. For example, while an animation of a circle bouncing in a very regular manner is likely to be interpreted as inanimate movement (a bouncing ball), calling the circle a 'prince' may encourage the viewer to find an interpretation of the movement that matches the animacy of the lexical item. In this example, one could come up with a story about a prince jumping on a trampoline (and some participants did). Similarly, a circle trying to climb up a slope might look very animate-like, but calling the circle a 'stone' may

cause the viewer to come up with an interpretation in which the stone was pushed with such force that it could withstand gravity and roll upwards. This makes it unclear whether the factor perceptual animacy really measured what it should measure, namely the impression of animacy people would get from purely perceptual features, i.e., movements. Experiment 2 was set up to deal with this potential complication.

## 3.4. Experiment 2

Experiment 2 was similar to Experiment 1, except that we replaced the lexical labels by nonsense words. These words were chosen in such a way to avoid intuitions about the animacy of the word as much as possible. In this way, we expected to minimize the chance that the lexical labels would influence the interpretation of the movements, and to get a clearer picture of the effects of perceptual animacy on referential choices.

### *3.4.1. Methods*

#### 3.4.1.1. Participants

Fifteen undergraduate students from Tilburg University participated in this study as speakers. Another 15 naive participants acted as addressees. Ten speakers and nine addressees participated for course credit; the others volunteered. All gave their consent to the use of their data. None of them had participated in Experiment 1.

#### 3.4.1.2. Materials

The animations were identical to those used in Experiment 1. However, instead of real words, the lexical labels consisted of nonsense words. These words were constructed by altering the real words from Experiment 1, while keeping the length in characters and the number of syllables the same (e.g., *daptinder* from *padvinder* ('boy scout')). Together with a number of real words, these constructed words were entered into a pretest in which nine participants indicated for each word whether they knew the word or not, and if not, to what degree they thought the word could refer to a person, an animal, or a thing. Participants marked their answers on five-point Likert scales (e.g., 'very likely a person' to 'very likely NOT a person'). Eight nonsense words that were indicated as 'unknown' by all participants, and had average scores around the middle of all three scales were selected for the present experiment. As in Experiment 1, they were presented together with the target figure just before the start of each

animation. To ensure that the labels would be interpreted as nouns referring to the target figures, the labels were preceded by the phrase *dit is een* 'this is a'. A list of all nonsense words can be found in Table A2 in Appendix A.

3.4.1.3. Procedure

The procedure was similar to that of Experiment 1, except that the addressee for the story retellings was no longer the experiment leader, but another naive participant. This was done to make the stories even more lively, which should reduce the number of missing data. To make the task more engaging, we also gave the addressees a task. For each object described by a nonsense word, they had to indicate on an answer sheet whether they thought this object was a person, an animal, or a thing. Because instructions were only given in written form, the speakers remained unaware of the nature of this task. Four speaker-addressee couples were tested in a face-to-face setting similar to that in Experiment 1; the other 11 couples communicated through Eye Catchers[6], because they were tested directly after another, unrelated, experiment that used this setup. Instructions were virtually identical to those of Experiment 1. After two practice trials, the experiment was started and the experiment leader left the room. It took about 20 minutes to complete the experiment.

3.4.1.4. Design

Since lexical animacy was held constant in this experiment by using nonsense words, the only independent variable was perceptual animacy. Participants were randomly assigned to one of two lists, which were created such that from a given item the perceptually animate version occurred on one list, and the perceptually inanimate version on the other. Items were presented in a random order. The same order was used in the two lists.

[6] See http://www.qconferencing.eu. Communication through Eye Catchers resembles face-to-face communication, because users can have direct eye contact (cf. Mol et al., 2011). Any effects of communication medium should not influence the results, since the use of Eye Catchers was counterbalanced across the animate and inanimate conditions. In addition, participants did see each other in person before the experiment and were experienced in using Eye-Catchers due to the unrelated experiment that directly preceded the presently discussed experiment. Removing the four participants that communicated face-to-face from the analysis did not change the results.

3.4.1.5. Data coding and statistical analyses

The coding scheme and the procedure for statistical analysis were identical to that of Experiment 1. We excluded six cases (5.0%) because either the transitive action was not described, or the given lexical label was not used.

*3.4.2. Results*

3.4.2.1. Choice of referent

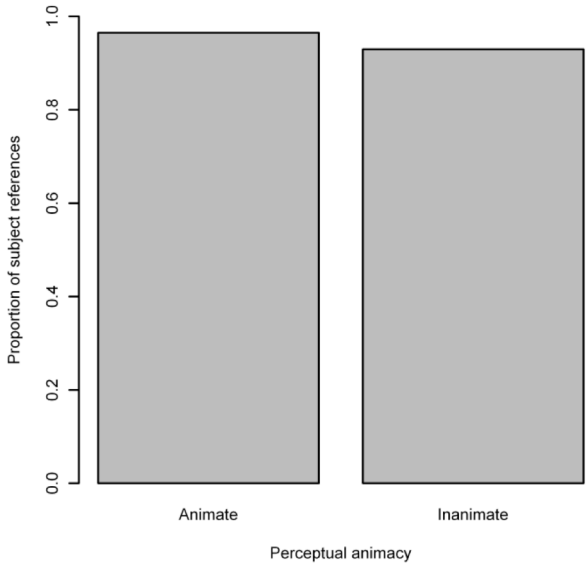


**Figure 3.7.** Proportion of subject references to the target figure in the description of the transitive action in Experiment 2, by its perceptual animacy.

Figure 3.7 shows the proportion of references to the target figure as the subject of the clause describing the transitive action as a function of perceptual animacy. Target referents were made the subject of the sentence in the majority of the cases (perceptually animate referents: 96.5%; perceptually inanimate referents: 93.0%). There was no significant effect of perceptual animacy, $\beta$ = 0.95, SE = 1.03, $p$ = .36. Random slopes were not included, since they did not improve model fit.[7]

[7] Although it was not the aim of this experiment to also investigate effects of discourse salience, we did perform separate analyses on the two types of context (discourse salient: referent was mentioned in the directly preceding sentence; n = 75, discourse non-salient: referent was not mentioned in the directly preceding sentence; n = 39). However, because the number of observations in each set was too small to

3.4.2.2. Choice of referring expression

Figure 3.8 shows the proportion of pronoun references to the target referent in the descriptions of the transitive action as a function of perceptual animacy. Pronouns were used more frequently when the target referent was perceptually animate (87.7%) than when it was perceptually inanimate (68.4%). The effect of perceptual animacy was significant, $\beta$ = 1.31, SE = 0.52, $p < .05$. Random slopes were not included, since they did not improve the model's fit.[8]

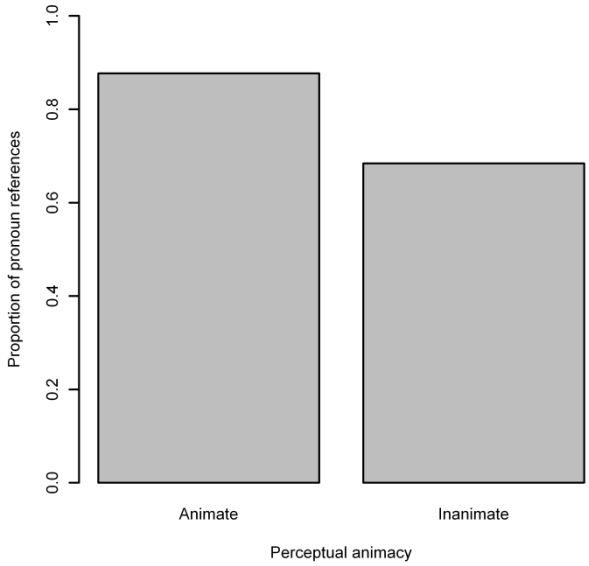


**Figure 3.8.** Proportion of pronoun references to the target figure in the description of the transitive action in Experiment 2, by its perceptual animacy.

support a mixed model with a full random effects structure, we performed two repeated measurement ANOVAs on arc-sine transformed proportions of subject references. These analyses showed no effects of perceptual animacy in either data set: discourse salient, $F1(1,14) = 1.00$, $p = .33$; $F2(1,7) = 1.00$, $p = .35$; discourse non-salient, $F1(1,10) = 0.51$, $p = .49$; $F2(1,6) = 0.57$, $p = .48$.

[8] Separate analyses were performed for discourse salient referents (n = 75) and discourse non-salient referents (n = 39). Repeated measurement ANOVAs (see footnote 7) showed a marginally significant effect of perceptual animacy in discourse salient contexts, $F1(1,14) = 4.53$, $p = .05$; $F2(1,7) = 3.80$, $p = .09$. In discourse non-salient contexts, the effect of perceptual animacy was significant but only over participants, $F1(1,10) = 7.34$, $p < .05$; $F2(1,6) = 0.20$, $p = .67$. In addition, we investigated whether the effect of animacy held for referring expressions in both subject and non-subject position. However, there were very few instances of non-subject expressions referring to the target figure (n = 6). When we omitted these cases from the analysis, the effect of perceptual animacy remained significant, $\beta = 1.09$, SE = 0.53, $p < .05$.

*3.4.3. Discussion*

The aim of Experiment 2 was to exclude the possibility that the perceived animacy of referents is influenced by the animacy of their lexical descriptions. By using nonsense words for which people have no strong intuitions about animacy, we controlled for this possible influence. The results largely confirmed the results of Experiment 1. An effect of perceptual animacy was found on the choice of referring expression in the descriptions of the transitive action. Here, perceptually animate referents were more likely to be pronominalized than perceptually inanimate referents. As before, perceptual animacy did not affect the choice of whether the referent was placed in subject position. These findings again suggest that perceptual animacy, a factor that is dependent on the visual context, can influence the choice of referring expression.

## 3.5. General discussion

This study was conducted to investigate whether perceptual animacy of referents influences referential choices, and how this interacts with lexical animacy. Firstly, the results of Experiment 1 confirm that lexical animacy affects the choice of referent for the subject position: Lexically animate referents were more likely to be mentioned as the subject of a transitive sentence. This is in line with research indicating that animacy affects grammatical function assignment (Bock et al., 1992; Dahl & Fraurud, 1996; McDonald et al., 1993). The effect was confined to contexts in which the referent was not discourse salient, suggesting that discourse salience may overrule effects of animacy (cf. Prat-Sala & Branigan, 2000). Perceptual animacy, on the other hand, did not seem to influence the choice of referent for the subject position, even when lexical cues for animacy were not present (Experiment 2). This suggests that visual cues can be overridden by lexical cues, which is not what would be predicted under the assumption that derived accessibility can override inherent accessibility.

Secondly, lexical animacy influenced the choice of referring expression: Also in line with previous research (Dahl & Fraurud, 1996; Fukumura & Van Gompel, 2011; Yamamoto, 1999), lexically animate referents were more frequently pronominalized, although this effect might be partly mediated by grammatical function. More importantly, the results of both Experiment 1 and Experiment 2 show that in addition to lexical animacy, perceptual animacy also affects pronominalization. This supports the hypothesis that both cues for animacy affect a referent's conceptual accessibility.

As predicted, perceptual animacy turned out to be a stronger factor in determining the choice of referring expression than lexical animacy, at least for discourse non-salient referents. This suggests that in this case visual cues can override lexical cues, in line with the idea that derived accessibility can override inherent accessibility. This is also consistent with Nieuwland & Van Berkum (2006), who found that contextual cues overruled lexical-semantic cues in comprehension. Surprisingly, in our study this does not only hold for lexically inanimate objects, which can be conceptualized as animate in a certain context, but also for lexically animate entities, which seem to become less animate when they move in an inanimate way.

While our materials were abstract, people readily conceptualized objects as more animate or more inanimate based on motion cues, as was already shown in the animacy pretest. That participants easily accommodated their conceptualizations of objects to the perceptual context, independently of their lexical animacy, can also be seen from the way the objects were described in Experiment 1. For example, lamps and snowflakes were described as 'happy' or 'afraid', and a handbag 'tries' or 'decides' to go in a certain direction. Conversely, there were some indications that lexically animate objects were conceptualized as inanimate. Besides the use of predicates that are associated (but not exclusively) with inanimate movement, such as 'fall', 'bounce', and 'slide', there were also a few cases in which objects were referred to with pronouns of which the gender did not match with the noun (e.g., using *hij* 'he', 'it' to refer to a queen). Although these cases might be errors, they were confined to the lexically animate–perceptually inanimate condition, suggesting that conceptual information associated with the lexical item was overruled by a different conceptualization. Still, given the gradual nature of animacy (Comrie, 1989), most objects were probably conceptualized as neither fully animate nor fully inanimate, both in the experiment and in the animacy pretest. The crucial point is, however, that manipulating motion alone made some objects appear more animate than others, and the same manipulation also affected pronoun use. Hence, objects that were statistically more likely to be conceptualized as animate were statistically more likely to be pronominalized.

These results have several theoretical implications. First of all, they suggest that which entity becomes the subject of the sentence is affected by the animacy of the lexical items, but not by the conceptualization of the entity based on perceptual cues. This seems inconsistent with a conceptual accessibility account of linearization (e.g., Bock & Warren, 1985; Prat-Sala & Branigan, 2000; Van Nice & Dietrich, 2003a).

According to such an account, it is the conceptual (rather than the lexical) accessibility of animate entities that causes them to be mentioned earlier in the sentence (either through a direct link to word order, or indirectly via grammatical function). Consequently, a lexically inanimate entity that is conceptualized as animate, such as a stone climbing up a slope, should also be more accessible according to this account.

However, others have linked conceptual accessibility to the retrieval of lexical items (Branigan et al., 2008; Levelt, Roelofs, & Meyer, 1999). According to Levelt et al. (1999), there is a close connection between concepts, lemmas and word forms. When a concept is highly accessible, this speeds up lemma retrieval, and, assuming incremental sentence production, the corresponding lexical item is produced first. This might explain the dominance of lexical animacy in referent choice. A crucial factor here might be the nature of the task. In our experiments, the lexical labels were presented before the start of the animations. Therefore, in Experiment 1 the concepts associated with the lexical items were already activated before any conceptualizations on the basis of perceptual cues were made. Hence, any contextually induced conceptualizations could have been overruled by the animacy of the activated lemmas when participants mentioned the referents. Our results might therefore have been different if we had presented the lexical labels after the animations.

In addition, Van Nice and Dietrich (2003b) found that when participants described pictures that remained in view, the animacy of the agent and the animacy of the patient both affected the rate of passivization, but there were no interactions between the two factors, suggesting that speakers were not actively comparing properties of different entities. However, when Van Nice and Dietrich had the participants describe the pictures from memory, they did find an interaction. They attributed this effect to 'compressed' processing of entities, such that entities would be processed almost simultaneously in this task, whereas they would have been processed serially when participants had more processing time. Similarly, they argued that the interaction between discourse salience and animacy found by Prat-Sala and Branigan (2000; i.e., context may overrule effects of (lexical) animacy) could be due to the fact that information about both factors was already present before participants started speaking, also leading to compressed processing. However, while in our experiments all information about the referents was also given before the participants started speaking, we did not find an interaction between target and competitor animacy (see footnote 2). In addition, while we did not present any linguistic context before the participants started speaking, we did find evidence that discourse context overruled

animacy effects. More research is needed to investigate how choices in language production are affected by the nature of the task.

A second implication of our findings is that the choice of referring expression can be influenced by non-linguistic, perceptual information. This is unexpected under an account in which the choice of referring expression is only determined by local discourse factors, such as givenness, grammatical function and topichood (e.g., Ariel, 1990; Grosz, Joshi, & Weinstein, 1995; Gundel, Hedberg, & Zacharski, 1993; Kehler, Kertz, Rohde, & Elman, 2008; Stevenson, 2002). As also found by Fukumura and Van Gompel (2011), the inherent accessibility of referents can influence the choice of referring expression. Our results extend this finding by showing that it is not just an effect of conceptual representations associated with lexical items. Rather, it involves a level of non-linguistic representation (e.g., Arnold, 2010; Bock & Warren, 1985), which may be shaped by perceptual properties, such as an object's movements. This also means that although animacy effects have been attributed to inherent accessibility, i.e., accessibility based on intrinsic properties of a (lexical) concept, these effects are also dependent on the (perceptual) context, and hence are partly driven by derived accessibility. It remains an open question whether such non-linguistic, perception-based representations should always be activated in linguistic tasks, or primarily play a role in cases of violations of canonical animacy such as in the present study. However, there is evidence that they are relevant also in more 'everyday' language use (see Rosenbach, 2008 for an overview). In addition, our finding that even lexically animate entities may be treated as less animate in reference when their movements appear less animate might suggest that our results are not just due to the participants' familiarity with cartoons or fairy tales, in which this kind of animacy shift is rare.

Combining our two main findings, a pattern emerges that seems contradictory: Whereas a referent's perceptual animacy can override lexical animacy in determining the choice of referring expression, this does not happen in the choice of referent for the subject position. This is not in line with accounts of referent accessibility in which the choice of referent for first mention and the choice of referring expression are both dependent on the conceptual accessibility of mental representations (e.g., Arnold, 2008, 2010). For example, in her Expectancy Hypothesis, Arnold (2008) proposes that accessibility is a catchall term for different factors that correlate with the probability that an entity will be mentioned again. However, the present results suggest that the different cues for animacy cannot easily be gathered under this single term.

Other researchers have suggested that choice of referring expression and choice of referent for first mention should be dissociated (e.g., Kehler et al., 2008; Stevenson, Crawley, & Kleinman, 1994). Typically, a distinction is made between bottom-up factors, such as grammatical function and information structure, and top-down factors, such as coherence, discourse topicality and general expectations about what will happen next. The choice of referring expression is assumed to be affected by bottom-up factors, similar to centering theory (Gordon, Grosz, & Gilliom, 1993; Grosz et al., 1995), while the likelihood that an entity will be mentioned next is affected by top-down factors. However, our finding that perceptual animacy influences the rate of pronominalization might also seem hard to reconcile with this account. Perceptual animacy can be regarded as a top-down factor, since it is dependent on context rather than linguistic properties. By analogy with other top-down factors such as coherence, perceptual animacy should be expected to affect likelihood of next mention, but not choice of referring expression.

An explanation for the apparent contradiction may lie in the task-dependent effects outlined above. While the choice of referent and the choice of referring expression may both be influenced by the referent's extra-linguistic conceptual accessibility, in the case of choosing a referent for the subject position this process may receive competition from the animacy of the lexical items when these are already given. That is, the presentation of a lexical item activates the corresponding lexical concept, which may boost the speed of retrieval of the lexical item in production (e.g., Bock & Irwin, 1980). This boost may be larger in the case of an animate concept. Hence, the quick retrieval of animate lexical items may make them more likely to be placed in subject position, even when the referent is conceptualized as inanimate due to the perceptual context. The choice of referring expression, on the other hand, is less dependent on the speed of lexical retrieval. Although it has been found that lexical information from the antecedent noun is activated when retrieving a pronoun (e.g., Meyer & Bock, 1999; Schmitt, Meyer, & Levelt, 1999), the choice of whether or not to use a pronoun in the first place can probably be made directly on the basis of the non-linguistic conceptual representation of an entity. Hence, this choice may be more driven by accessibility derived from perceptual cues than the choice of referent.

Another issue is that since we investigated the number of subject references, and not first mention per se, we cannot exclude the possibility that perceptual animacy does have an effect on linearization independently of grammatical function, as has been shown for lexical animacy (e.g., Branigan & Feleki, 1999; Branigan et al., 2008; Prat-

Sala & Branigan, 2000). Whereas Dutch, the language of our experiments, has a relatively free word order, and thus allows for the investigation of linearization independently of grammatical function, our data were not structured enough to analyze this. For example, as we were dealing with relatively spontaneous speech, most of our selected clauses were not clear-cut sentences, which makes it hard to determine where exactly sentences begin. Therefore, we did not further pursue these analyses, and we leave this issue to future research.

Still, it is not clear how our finding that perceptual animacy affects rate of pronominalization would be accounted for in models of language production. It could be argued that this effect is mediated by other factors, such as information structure or agency. For example, animate entities may be more likely to be topics (e.g., Givón, 1983). Hence, it might be the case that perceptually animate entities are more likely to be pronominalized because they are topics. This explanation would be in line with a centering-type account. However, this is not a likely explanation. Firstly, we have seen that the effects of perceptual animacy remained intact when we controlled for discourse salience. Although entities that had been mentioned in the previous clause were likely to be pronominalized, entities that had not been mentioned (and thus were not likely topics) were more likely to be pronominalized when they were perceptually animate. Secondly, the target referent was always presented just before each animation, and it was also the first entity to appear in the screen at the start of each animation. We assume that this made all target referents equally likely candidates for the discourse topic across the conditions.

As for agency, this is another factor that is likely to affect a referent's conceptual accessibility (e.g., Van Nice & Dietrich, 2003b), but it is hard to disentangle it from animacy, since agents are often animate. In our experiments, we kept the agency of the objects constant in the sense that it was always the target figure that moved in the episode of interest (the transitive action), while the competitors remained still. Thus, although the mere fact that the objects moved might already have increased their conceptual accessibility, this was at least the same for both the animate and the inanimate moving objects. However, a valid argument against this is that agency is itself a gradient notion (e.g., Dowty, 1991). That is, the inanimate movements in our experiments may have been inherently less agentive than the animate movements. Dowty (1991) proposed an entailment hierarchy of agentivity, in which entities are considered more agentive the more properties of a prototypical agent they possess. The hierarchy is headed by the proto-agent properties volition (i.e., having a will) and

sentience (i.e., being conscious, being able to perceive), both of which entail animacy. On this view, we cannot exclude the possibility that our perceptually animate objects were also more prototypical agents than our perceptually inanimate objects. Since it would be difficult to completely disentangle perceptual animacy from perceptual agency by using motion alone, future studies could find different ways of manipulating perceptual animacy to tease these two factors apart.

These issues notwithstanding, the present study has shown that referential choices in Dutch can be influenced by factors that go beyond the linguistic context, and may even be perceptual in nature. We have focused on the interaction between two sources of animacy, and although many studies have shown that animacy is not reducible to an epiphenomenon of some other accessibility-related factor, it is clear that it interacts with many other factors (e.g., Comrie, 1989; Rosenbach, 2008; Van Nice & Dietrich, 2003a). Future research should investigate more interactions between accessibility-related factors in referential choices, especially between linguistic and perceptual factors, and between discourse-related and referent-intrinsic factors, possibly also using more online measures such as eye movements.

In summary, the present study provides evidence for differential effects of perceptual and lexical cues for animacy on referential choices. Perceptual animacy appeared to overrule lexical animacy in the choice of referring expression, extending previous findings. On the other hand, the choice of referent for the subject position appeared to be affected only by lexical animacy. The results raise new questions about the nature of animacy effects on referential choices in particular and of accessibility in general.

# Chapter 4

# Animacy in Belgian and Netherlandic Dutch

**Abstract**

It has been argued that animate entities tend to be referred to with more attenuated expressions than inanimate entities, because they are more accessible in memory. Two previously untested claims made for Dutch suggest that the situation may be more complex. Firstly, it has been stated that full pronouns can only refer to animate entities, while reduced pronouns can also refer to inanimate entities. This seems to be inconsistent with the accessibility account. Secondly, inanimate entities may be pronominalized less to avoid gender-marked expressions when grammatical gender is unclear. Using a sentence completion task, we tested these claims by investigating the effect of animacy both on the choice of pronouns versus more specific expressions and on the choice of reduced versus full pronouns in Dutch. We compared speakers of Netherlandic Dutch, who have generally lost intuitions about masculine and feminine grammatical gender of nouns, with speakers of Belgian Dutch, for whom these intuitions are still relatively intact. Both groups showed an effect of animacy on pronominalization, suggesting that this effect cannot only be explained as a gender avoidance strategy. At the same time, the choice between full and reduced pronouns cannot be explained by accessibility. Implications for theories of reference production are discussed.

**This chapter is based on:**

Vogels, J., Maes, A. A., & Krahmer, E. J. (In press). Choosing referring expressions in Belgian and Netherlandic Dutch: effects of animacy. Accepted for publication in *Lingua*.

### 4.1. Introduction

Among the many types of expressions that can be used to refer to an entity, pronouns encode the least information about the referent (leaving aside morphologically invisible ways of referring, such as the use of zero anaphora). Languages differ in the amount of information encoded in pronouns, but for the Germanic languages it generally holds that personal pronouns convey information about person, number, gender, and/or case. Given the general nature of these features, pronouns can be highly ambiguous in everyday language: They can refer to many different things. Therefore, the restricted informational content of pronouns requires that their referents somehow be salient, i.e., be in the focus of attention, to be correctly identified (e.g., Ariel, 1990; Givón, 1976; Gundel, Hedberg, & Zacharski, 1993). Several discourse factors have been found to influence the salience of a referent, for instance whether it was the subject, the topic, or the first mentioned entity in the previous sentence (e.g., Gernsbacher & Hargreaves, 1988; Gordon, Grosz, & Gilliom, 1993; Grosz, Joshi, & Weinstein, 1995). In addition, the salience of a referent can also be influenced by its intrinsic properties, such as its animacy. Human referents may be inherently more salient than other animate referents (e.g., animals), which are in turn more salient than inanimate referents. The animacy hierarchy, given in (1), which was originally proposed as an implicational hierarchy to explain certain grammatical phenomena (e.g., Comrie, 1989), may therefore also be seen as a salience hierarchy.

(1) *Animacy hierarchy*
Human > Animate > Inanimate

Indeed, many psycholinguistic studies have shown that human or animate referents are more likely to be placed in subject position, or to be mentioned earlier in the sentence, than inanimate referents (e.g., Bock, Loebell, & Morey, 1992; Bresnan, Cueni, Nikitina, & Baayen, 2007; Prat-Sala & Branigan, 2000; Rosenbach, 2005; Van Nice & Dietrich, 2003a). These effects are generally explained as conceptual accessibility effects. That is, mental representations of human or animate entities are more activated in memory and therefore more easily retrieved than those of inanimate entities (Bock & Warren, 1985). Crucially, if animacy affects the accessibility of a referent in memory, it is also expected to affect what type of referring expression is chosen to refer to that referent. It is generally assumed that more accessible referents

are referred to with more attenuated referring expressions (e.g., Ariel, 1990). According to Ariel, referents with a low accessibility are preferred to be marked by a full definite description; referents of intermediate accessibility can be marked by demonstrative pronouns, for example; and highly accessible referents are preferably marked by unstressed, reduced pronouns or zero anaphora. This is expressed by the accessibility scale in (2). Mapping the animacy hierarchy in (1) on the accessibility scale in (2) results in the prediction that the more animate a conceptual representation of a referent is, the more attenuated the expression referring to that referent will be.

(2) *Accessibility scale* (simplified version of the scale presented in Ariel, 1990)
Zero > Reduced pronoun > Full pronoun > Proximate demonstrative > Distal demonstrative > Full NP

Indeed, in a corpus of Swedish texts, Dahl and Fraurud (1996) found that antecedents denoting humans were more than four times as likely to be referred to with pronouns as inanimate antecedents (as opposed to full noun phrases; see also Yamamoto, 1999 for corpus data on English and Japanese). In a story completion experiment, Fukumura and Van Gompel (2011) found a similar (but weaker) tendency in the frequency of the English third person plural pronoun *they*, which was higher when participants referred to animate entities than when they referred to inanimate entities. Fukumura and Van Gompel argued that this effect was not due to the avoidance of the ambiguity that is inherent to pronouns, since the same effect was found in contexts in which a pronoun was not ambiguous. In addition, they found that the effect was independent of the grammatical function of the antecedent noun, suggesting that it cannot be explained as an effect of discourse salience. Fukumura and Van Gompel proposed that references to animate entities are more often attenuated than references to inanimate entities because human or animate referents are inherently more accessible in memory. Therefore, less information from the conceptual representation needs to be encoded in a referring expression to activate that representation.

The direct mapping of the animacy of the referent to the degree of attenuation in referring expressions is not unproblematic, however. Firstly, it has been claimed that in languages that have a distinct set of pronouns that are phonologically or morphologically reduced and cannot be stressed, these reduced forms can be used to refer to both animate and inanimate entities, while their full counterparts, which can

be stressed, can only refer to animate (human) entities (e.g., Cardinaletti & Starke, 1996). This goes exactly counter to the idea that expressions referring to animates are more likely to be attenuated, since that would predict that reduced pronouns are more frequent in reference to animate entities than full pronouns. Secondly, in both English and Swedish, the languages investigated in the studies mentioned above, the relation between the animacy of the referent and the choice of a certain referring expression has been grammaticalized to some degree. For example, masculine and feminine pronouns in those two languages are used almost exclusively for animate (human) referents (e.g., Corbett, 1991). In English, the neuter pronoun *it* is used to refer to inanimate entities, and, when used anaphorically, English demonstrative pronouns, which do not have a gender distinction, often have antecedents that are not individuals (e.g. propositions or composite entities; Brown-Schmidt, Byron, & Tanenhaus, 2005). In Swedish, the common gender pronoun *den* and the neuter pronoun *det* typically refer to non-human entities, as illustrated by the ungrammaticality of (3b) as opposed to (3a).

(3) a. *Jag gick på Skyfall istället för på the Hobbit$_i$, eftersom jag tyckte*
I went to S instead of to TH because I thought
*att den$_i$ verkade tråkig.*
that that/it seemed boring
'I went to see Skyfall instead of The Hobbit, because that one seemed boring.'

b. **Jag gick till Simon istället för till Hugo$_i$, eftersom jag tyckte att*
I went to S instead of to H because I thought that
*den$_i$ verkade tråkig.*
that/it seemed boring

Hence, the tendency to use different types of referring expressions to refer to animate and inanimate entities may be influenced by the fact that some forms already encode animacy. This raises the question whether the effect of animacy on the choice of referring expression can be generalized to languages that do not have animacy grammaticalized in the pronominal system.

In this chapter, we investigate whether and how animacy affects the choice of referring expression in two varieties of standard Dutch. Dutch is a language well-suited for investigating the two issues raised above. Firstly, it has a set of both full and reduced third person pronouns, which allows us to test whether animate entities are

referred to with more attenuated expressions also on a more fine-grained level, or whether there is an opposite tendency within the use of pronouns, as suggested by Cardinaletti and Starke (1996). Secondly, Dutch has no pronouns that exclusively refer to animate or inanimate entities. In principle, masculine, feminine and neuter personal pronouns can all be used to refer to animates as well as inanimates, as in German.[1] The same holds for demonstrative pronouns, as illustrated by the Dutch variants of (3), given in (4).

(4) a. *Ik ben naar Skyfall gegaan in plaats van naar The Hobbit$_i$, omdat die$_i$ me saai leek.*
I am to S gone in place of to TH because that meboring seemed
'I went to see Skyfall instead of The Hobbit, because that one seemed boring.'

b. *Ik ben naar Simon gegaan in plaats van naar Hugo$_i$, omdat die$_i$ me saai leek.*
I am to S gone in place of to H because that me boring seemed
'I went to see Simon instead of Hugo, because he seemed boring.'

Thus, whereas speakers of English and Swedish need to take into account the animacy/humanness of the referent to choose a pronominal referring expression, speakers of Dutch only need to select a pronoun with the correct gender. If personal pronouns are more frequent in reference to animates than to inanimates also in Dutch, this would provide additional evidence for the hypothesis that the increased conceptual accessibility of animate entities makes them more likely to be referred to with attenuated expressions.

Another reason why Dutch is well-suited for investigating the questions at hand is that there are two varieties of the Dutch standard language, spoken in the Netherlands and Belgium, respectively, that differ in the degree to which speakers have intuitions about the grammatical gender of nouns. While Dutch originally had a

[1] In references to humans, it is more common in Dutch to select a pronoun based on the semantic gender of the noun, while in German the grammatical gender might also be used (although this may be less common in colloquial German). For example, the German neuter pronoun *es* may be used when the antecedent is a grammatically neuter but semantically feminine noun such as *Mädchen* 'girl' (Dahl, 2000), which would be odd in Dutch.

three-way gender distinction (masculine, feminine, and neuter) in the nouns, modern standard Dutch nouns morphologically distinguish only between common gender (*de*-nouns; e.g., *de man* 'the man', *de vrouw* 'the woman') and neuter gender (*het*-nouns; e.g., *het kind* 'the child'). On the one hand, most speakers of Dutch in the Netherlands do not intuitively know the original (masculine or feminine) gender of inanimate common gender nouns. According to Audring (2006; 2009), these speakers therefore exhibit a tendency to avoid the choice between masculine and feminine pronouns when referring to inanimate antecedents. Instead, they switch to using demonstrative pronouns or definite descriptions, which are not marked for masculine or feminine gender (demonstratives do distinguish between common and neuter gender forms). A similar tendency has been found for Swedish, where speakers sometimes use the common gender pronoun *den* to refer to persons when the antecedent's natural gender is unknown (Josefsson, 2010). This gender avoidance strategy may be another explanation why inanimate entities are less frequently pronominalized than animates.

In the variety of Dutch spoken in Belgium, on the other hand, many speakers still have intuitions about whether a *de*-noun is masculine or feminine. Therefore, it is unlikely that gender mismatches are a reason for these speakers to avoid pronouns when they refer to inanimate entities. Comparing the two varieties of Dutch thus allows us to test whether possible effects of animacy on the choice of referring expressions in Dutch can be explained by a general conceptual accessibility effect, or whether a gender avoidance strategy also plays a role, which has not been empirically investigated before. If the former is the case, an effect of animacy should be present in both varieties. If avoidance plays a role, the animacy effect should be larger for speakers of Dutch from the Netherlands than for speakers of Dutch from Belgium. Before moving on to the present study, which was conducted to test these predictions, the next section first presents an overview of the Dutch pronominal and nominal gender systems.

### 4.2. Pronouns and grammatical gender in Dutch

The Dutch third person personal pronouns are similar to the English ones in that there are masculine, feminine and neuter forms in the singular, and no gender distinctions in the plural. They differ from those in English in that almost all personal pronouns, both subject and object forms, have morphologically distinct full and reduced variants.

In addition, masculine and feminine pronouns are not restricted to animate reference. Table 4.1 lists all third person subject pronouns used in the standard language.[2]

**Table 4.1.** The gender system of Dutch third person subject pronouns

| | masculine | feminine | neuter | plural |
|---|---|---|---|---|
| full pronouns | *hij* | *zij* | - | *zij* |
| reduced pronouns | *ie*[a], *die* | *ze* | *het*, '*t*[b] | *ze* |

[a] The reduced pronoun *ie* cannot occur sentence-initially.
[b] The neuter pronoun *het* 'it' is categorized here as reduced pronoun, as it patterns syntactically with the reduced forms (Coppen, Haeseryn, & De Vriend, 2002).

It has been argued that the reduced pronouns cannot be stressed, coordinated or modified, while the full pronouns can (e.g., Cardinaletti & Starke, 1996; Coppen Haeseryn, & De Vriend, 2002). In addition, according to the Dutch reference grammar (*Algemene Nederlandse Spraakkunst* (ANS); Coppen et al., 2002: §5.2.7), full pronouns (also in the object forms) typically refer to animate entities (unstressed *hij* 'he' being an exception), while reduced pronouns can be used to refer to both animate and inanimate entities. Hence, the sentence in (5a) is presented as infelicitous when *zij* 'they' refers to books, while (5b), with the reduced form *ze*, is fine in this context.

(5) a. # *Zij staan daar al een hele tijd.*
b. *Ze staan daar al een hele tijd.*
they stand there already a whole time
'They have been standing there for a long time.' (they = books)

While this constraint might not be as strict as presented here (see e.g., Van Bergen, Stoop, Vogels, & De Hoop, 2011), the same observation has been made for other languages, such as German and Italian, in which full or non-cliticized pronouns can only refer to humans (Cardinaletti & Starke, 1996). This does not seem to be in accordance with the accessibility scale, which suggests a strong tendency for more accessible referents to be referred to with more attenuated expressions (Ariel, 1990; cf. Kaiser, 2011; Kaiser & Trueswell, 2008). Since reduced pronouns are more attenuated

[2] Note that the third person plural pronoun (*zij*, *ze*) is homophonous with the third person singular feminine pronoun. In addition, the reduced pronoun *die* 'he' is homophonous with the demonstrative pronoun *die* 'that'. However, the former cannot be stressed, while the latter can be either stressed or unstressed.

than their full counterparts, they should be more likely to refer to more accessible, and hence more animate, referents. To our knowledge, this has not yet been tested experimentally.

With respect to gender, while Dutch has masculine, feminine and neuter third person singular personal pronouns, Dutch demonstrative pronouns and definite noun phrases morphologically distinguish only two genders: common and neuter. This is shown in Table 4.2.

**Table 4.2.** The gender system of Dutch demonstrative pronouns and definite noun phrases

| | common | neuter | plural |
|---|---|---|---|
| demonstrative pronouns | *die, deze* | *dat, dit* | *die, deze* |
| definite noun phrases | *de man, de vrouw* | *het kind* | *de mannen, de vrouwen, de kinderen* |

As Old English and Old Swedish, Old Dutch (approx. 500-1150) had a fully marked three-way nominal gender system (e.g., Audring, 2009). Gender distinctions began to erode with the simplification of declension systems in the Middle Dutch period, until at some point the original masculine and feminine genders collapsed into one common gender category (Geerts, 1966), resulting in only two nominal genders in the present-day standard language: common gender nouns that take the definite article *de* (*de*-nouns), and neuter gender nouns that take the definite article *het* (*het*-nouns). However, some dialects (among which the southern varieties Flemish, Brabantian and Limburgish) retain the three-way gender distinction until the present day, marked morphologically by different forms of articles and adjectives (e.g., Sint Niklaas dialect: *ne grote man* 'a tall man' (masc.), *een grote vrouw* 'a tall woman' (fem.), *e groot kind* 'a tall child' (neut.); De Vogelaer & De Sutter, 2011). This has resulted in a situation in which the distinction between masculine and feminine nouns is no longer felt in the variety of standard Dutch spoken in most of the Netherlands (excluding south-eastern dialects; henceforth Netherlandic Dutch),[3] while this distinction is still in use in the variety of standard Dutch spoken in Belgium (Flanders; henceforth Belgian Dutch;

[3] There are still certain morphological criteria that indicate the original grammatical gender of an inanimate common gender noun. For example, words ending in *–ing* or *–heid* are always feminine. However, even these criteria are not part of the internal grammar of many speakers, given frequent agreement errors such as *de regering en zijn ministers* 'the government$_{F}$ and its$_{M/N}$ ministers'.

Coppen et al., 2002), probably because of the strong presence of dialects that still have a three-way gender system (De Vogelaer, 2006).

In sum, there is a mismatch between nominal and pronominal gender in Netherlandic Dutch, with three pronominal genders (masculine, feminine, and neuter), but only two nominal genders (common and neuter). Since Dutch does not have a common gender (or 'uter') personal pronoun, such as Swedish *den*,[4] common gender nouns have to be pronominalized using either a masculine (*hij*, *ie*) or a feminine (*zij*, *ze*) pronoun, neither of which matches the grammatical gender of the noun. In reference to persons, speakers can use the referent's natural (i.e., biological) gender to guide their choice of a pronoun (e.g., Dahl, 2000). In reference to non-persons, however, this is not possible. Here, speakers of Netherlandic Dutch typically use a masculine pronoun, which can be considered unmarked with respect to the feminine forms (De Vogelaer, 2009). According to Audring (2006; 2009), the choice for a pronoun in Netherlandic Dutch is becoming more semantically driven. That is, a process is taking place in which the choice for a masculine, feminine or neuter pronoun is determined more by semantic properties of the referent than by (arbitrary) grammatical gender (see also Corbett, 1991). Audring proposes that the relevant semantic property is the degree of individuation of the referent: The masculine pronouns *hij* and *ie* mark a relatively high degree of individuation (e.g., animals or concrete inanimate objects),[5] while the neuter pronoun *het* marks low individuation (e.g., mass nouns).[6]

For speakers of Belgian Dutch, there is no mismatch between nominal and pronominal gender, and hence they can choose a pronoun based on the original grammatical gender of the antecedent. Thus, masculine pronouns (*hij* or *ie*) typically only refer to originally masculine nouns, feminine pronouns (*zij* or *ze*) to originally feminine nouns, and neuter pronouns (*het*) to neuter nouns. There is some evidence that the three-way gender system is gradually eroding in Belgian Dutch as well, probably due to the decline in the use of dialects in favor of the standard language (De Vogelaer, 2006; De Vogelaer & De Sutter, 2011). Especially younger speakers, who

[4] A reviewer notes that the development of Dutch demonstrative pronoun *die* may parallel that of Swedish *den*, in that *die* is on its way to become a personal pronoun (see also Audring, 2009, p. 104). Since this development is arguably still in an early stage, however, we would like to maintain the distinction between demonstrative and personal pronouns here. See also footnote 10.

[5] Even when an animal's natural gender is clearly female, such as in the case of *koe* 'cow', some speakers will use *hij* 'he' (Audring, 2009).

[6] The feminine possessive or object pronoun *haar* is sometimes used to refer to collectives or abstract nouns, even when these nouns were originally masculine or neuter (see Audring, 2006, footnote 9; Van der Sijs, 2004).

have less knowledge of the dialects, may be converting to a two-way system. However, this development is still in an early stage compared to Netherlandic Dutch, as witnessed by the frequent use of grammatically agreeing pronouns in Belgian Dutch (De Vogelaer & De Sutter, 2011). The difference between the two varieties is illustrated in (6), which is taken from the experimental items of the study presented here. In this example, the common gender noun *tent* 'tent', which takes the article *de*, was originally feminine, and is assumed to be still felt as such in Belgian Dutch. Thus, speakers of Belgian Dutch are expected to be more likely to use a feminine pronoun, as in (6b), while speakers of Netherlandic Dutch will be more likely to use a masculine pronoun (the default) to refer to an inanimate antecedent, as in (6a).

(6) a. Netherlandic Dutch

*Met slecht weer wordt de tent beschut. Is* ***hij / ie*** *gemakkelijk*
with bad weather becomes the tent$_{C}$ sheltered is he easily
*verplaatsbaar?*
movable

b. Belgian Dutch

*Met slecht weer wordt de tent beschut. Is* ***ze*** *gemakkelijk*
with bad weather becomes the tent$_{F}$ sheltered is she easily
*verplaatsbaar?*
movable
'In bad weather, the tent is sheltered. Can it be moved easily?'[7]

Crucially, however, speakers of Netherlandic Dutch may be uncomfortable with the use of a masculine pronoun here, since that would suggest that it refers to a masculine antecedent, while there may still be some awareness that the antecedent might possibly be feminine. Thus, the loss of intuitions about the masculine-feminine gender distinction could result in uncertainty about pronoun choice. It has been claimed that because of this uncertainty, there is a tendency in Netherlandic Dutch to avoid pronoun references to inanimate entities altogether (Audring, 2009). This leads to the use of other types of referring expressions, such as demonstrative pronouns (which do

[7] In this example and elsewhere, M stands for masculine gender, F for feminine gender, C for common gender, and N for neuter gender.

have common gender forms) or full noun phrases. Audring (2009) cites the following example:

(7) *De mummie zal eerst een CT-scan ondergaan voordat deze*
thec mummyc will first a CT-scanc undergo before this.onec
*tentoongesteld wordt in het Sakkara museum.*
exhibited becomes in theSakkara museum
'The mummy will first undergo a CT-scan before it is exhibited in the Sakkara museum.'
(Audring, 2009, p. 47)

In (7), the common gender proximal demonstrative *deze* 'this' is used to refer to the common gender noun *mummie* 'mummy', despite the fact that there is a preference for demonstratives to refer to the most recent antecedent rather than to the subject of the previous clause, and hence *CT-scan* would be a more preferred antecedent. If this use of a demonstrative pronoun is due to uncertainty about whether *mummie* should be referred to with a masculine or a feminine pronoun, the choice of referring expression should be less problematic for a speaker of Belgian Dutch who knows that *mummie* is grammatically feminine, and hence prefers the use of the feminine pronoun *ze*.

### 4.3. Predictions and experimental design

In this chapter, we investigate to what extent the observations about pronoun use in Dutch outlined in the previous section are in line with a conceptual accessibility account. This account assumes that animate entities are more accessible than inanimate entities, and hence that references to animates should be more reduced than references to inanimates. Firstly, this account predicts that reduced pronouns in Dutch (*ie*, *ze*) are more frequent than full pronouns (*hij*, *zij*) in reference to animate entities. Secondly, animate entities are predicted to be more likely to be pronominalized than inanimate entities. This effect should be present both in Netherlandic Dutch and in Belgian Dutch. According to the *gender avoidance account*, however, the animacy effect could also partly be explained by speakers being uncertain about choosing gender-marked pronouns for inanimate nouns. Hence, the tendency to use fewer pronouns for inanimate referents should at least be smaller in Belgian Dutch, since speakers of

Belgian Dutch do not need to avoid gender-marked forms due to the one-to-one mapping between nominal and pronominal gender. To test these predictions, we conducted a spoken sentence completion experiment with a group of Netherlandic Dutch speakers and a group of Belgian Dutch speakers, in which participants read sentences aloud and produced referring expressions for animate and inanimate antecedents.

## 4.4. Methods

### *4.4.1. Participants*

Twenty-four native speakers of Dutch from Belgium (14 female; aged 45-93; mean age 58.3), and seventeen native speakers of Dutch from the Netherlands (9 female; aged 27-65; mean age 48.2) participated in this study. The participants from the Netherlands did not speak any dialect with a three-way gender distinction (no Brabantian, Limburgish, Achterhoeks or Twents). The participants from Belgium were all from Brabantian or Flemish (East Flemish, West Flemish[8]) speaking areas. No participants from Belgium under the age of 40 were recruited, because of the possibility that knowledge of grammatical gender is declining in younger generations.

### *4.4.2. Materials and Design*

The materials included 16 short texts, each consisting of two sentences written in Standard Dutch. The first sentence always described a transitive event involving an animate and an inanimate entity, e.g., *Plotseling valt de inbreekster tegen de klok aan* 'Suddenly the (female) burglar falls against the clock'. The second sentence was always a question addressing the first sentence, in which the subject constituent was replaced by a gap, e.g., *Is ___ kapot?* 'Is ___ broken?'. Each text occurred in four different versions: (a) the subject of the question referred to the inanimate entity, which was the subject of the first sentence; (b) the subject of the question referred to the animate entity, which was the object of the first sentence; (c) the subject of the question referred to the inanimate entity, which was the object of the first sentence; (d) the subject of the question referred to the animate entity, which was the subject of the first sentence. The four versions are exemplified in (9).

[8] A tendency towards a (semantically motivated) two-way nominal gender distinction may also be seen in West Flemish (De Vogelaer & De Sutter, 2011). However, we did not observe such a trend in our two participants from West Flanders (including a 93 year old).

(9) a. *Plotseling valt de klok tegen de inbreekster aan. Is _____ kapot?*
'Suddenly the clock falls against the (female) burglar. Is _____ broken?'
b. *Plotseling valt de klok tegen de inbreekster aan. Is _____ nu verraden?*
'Suddenly the clock falls against the (female) burglar. Has _____ now been betrayed?'
c. *Plotseling valt de inbreekster tegen de klok aan. Is _____ kapot?*
'Suddenly the (female) burglar falls against the clock. Is _____ broken?'
d. *Plotseling valt de inbreekster tegen de klok aan. Is _____ nu verraden?*
'Suddenly the (female) burglar falls against the clock. Has _____ now been betrayed?'

Which entity was the target referent was made clear by a disambiguating context in the second sentence. For example, *Is ___ kapot?* 'Is ___ broken' can only refer to an inanimate entity, while *Is ___ nu verraden?* 'Has ___ now been betrayed?' can only refer to an animate entity. Hence, we expected participants to refer to the clock in (9a) and to the burglar in (9b). Versions (c) and (d) were included to control for the effect of grammatical function on the choice of referring expression. We used interrogative sentences with subject-verb inversion to enable the use of the masculine reduced pronoun *ie*, which cannot appear sentence-initially. The animate nouns had either feminine or masculine gender (corresponding to female and male natural gender, respectively), while the inanimate nouns always had feminine gender, according to the Van Dale dictionary.[9]

To check whether the nouns were felt as feminine by Belgian Dutch speakers, but not by Netherlandic Dutch speakers, 11 Belgian Dutch and 14 Netherlandic Dutch speakers not participating in the experiment were asked to make a forced choice between a feminine pronoun (*ze* 'she') and a masculine pronoun (*ie* 'he') to refer to a set of (originally feminine) inanimate nouns. This test consisted of 57 short texts, similar to the ones used in the experiment, except that the initial sentence only contained one noun phrase, which was always inanimate and feminine. There was only one version of each item. An example is given in (10).

[9] Until recently, most Dutch dictionaries gave the original gender of a lexeme, with the 'new' gender as felt by most Netherlandic Dutch speakers given in brackets. The latest editions often omit gender information when a lexeme is feminine for Belgian Dutch speakers but common (masculine) gender for Netherlandic Dutch speakers.

(10) *Weer valt de klok. Is _____ kapot?*
'The clock falls down again. Is ____ broken?'

Only nouns for which at least 75% of the Belgian Dutch speakers chose *ze* and at least 75% of the Netherlandic Dutch speakers chose *ie* were used in the experiment.

The materials were further checked for acceptability by 27 native speakers of (Netherlandic) Dutch, who were asked to indicate the degree of semantic well-formedness of the first sentence and the degree of coherence between the first sentence and the second sentence (with the gap filled in). Only materials with medium to high acceptability scores (e.g., 3 or above on a 5-point Likert scale) were selected, although we were not always able to avoid the use of slightly contrived sentences due to the structure of the materials. In addition, care was taken that the materials contained no words or expressions that are exclusively used in the Netherlands or Belgium, which was checked by the author of this dissertation (native Netherlandic Dutch) and his second supervisor (native Belgian Dutch). A full list of the experimental items can be found in Appendix B. In addition to the experimental materials, 20 fillers were constructed. These were similar to the experimental materials, but some had two animate or two inanimate entities (either feminine or masculine), and not all contexts were disambiguating. This was done to avoid task strategies and to encourage variation in the choice of referring expressions. The filler items had only one version.

The design of the experiment consisted of the two within-participant factors animacy (animate referent or inanimate referent) and grammatical function (subject or object), and the between-participant factor variety (Belgian Dutch or Netherlandic Dutch). This resulted in a 2 x 2 x 2 mixed design. Two stimulus lists were created, each of which contained two versions of each item: a version with the inanimate entity as the subject and its counterpart with subject and object swapped around (e.g., (9a) and (9c), or (9b) and (9d)). This was done because grammatical function was only included as a control, and it would have complicated the design if subject and object references were also distributed over multiple stimulus lists. Participants were randomly assigned to each list. The lists were divided into two blocks, to keep the two versions of the same item as far apart as possible. The filler items were intermixed pseudo-randomly with the experimental items, with at most two experimental items appearing in consecutive order.

*4.4.3. Procedure*

The experiment was run from a web server on a laptop using WWStim (Veenker, 2003). The items were presented one by one, and each sentence pair immediately appeared in its entirety upon mouse click. Participants were instructed to first carefully read the complete sentence pair silently, think of a word or phrase to fill in the gap, and then read the whole text aloud while filling in the gap. Participants were further instructed not to ponder too long, but there were no time restrictions on the trials. The experiment started with three practice trials, and it took about 10 minutes to complete the experiment. All speech was recorded.

To investigate the degree to which the Belgian participants would use the grammatical gender of nouns in choosing a pronoun when using the standard language, these participants were given a written gender test after the experiment, in which they had to make a forced choice between a masculine (*ie*) and a feminine (*ze*) pronoun to refer to originally feminine inanimate nouns (the same test as the material pretest). As a control, 5 participants from the Netherlandic Dutch group also performed this task. The test was done post hoc to avoid influences on the experiment. The results showed that across all Belgian participants, a feminine pronoun was used in 75% of the cases. Individual scores ranged from 39% to 95%. The scores correlated with age: Older participants were more likely to have higher proportions of feminine pronouns (Spearman's $\varrho = 0.61$; $p < .01$). Individual chi-square tests showed that for 7 participants, the number of feminine pronouns was not significantly ($p > .05$) higher than chance (but none of them used significantly more masculine than feminine pronouns). Because our main interest was in comparing groups of participants that differed in their intuitions about the three-way gender system, these participants were excluded from further analysis. The remaining 17 participants used feminine pronouns in 85% of the cases on average. All subsequent analyses are done using this subset of the Belgian participants.

The Dutch participants that completed the same task used a feminine pronoun in 29% of the cases, with a range from 0% to 44%. Individual chi-square tests showed that the scores were either not different from chance, or that there were significantly more masculine pronouns. A logistic regression analysis with language variety as a factor and the log odds for a feminine pronoun as the dependent variable showed that the Belgian participants used significantly more feminine pronouns to refer to originally feminine inanimate nouns than the Dutch participants ($\beta = 2.00$; SE = 0.15; $p < .001$).

*4.4.4. Data coding*

From the 17 Netherlandic Dutch participants (henceforth NLD speakers) and the remaining 17 Belgian Dutch participants (henceforth BD speakers), we transcribed all referring expressions used to fill in the gaps in the experimental items. We coded for the following categories: full noun phrases preceded by a demonstrative pronoun (e.g., *deze tent* 'this tent'); full noun phrases preceded by a definite article (e.g., *de koopman* 'the merchant'); demonstrative pronouns (*die*, *deze*, *dat,* or *dit*); masculine full pronouns (*hij*); feminine full pronouns (*zij*); masculine reduced pronouns (*ie*); feminine reduced pronouns (*ze*); neuter pronouns (*het*). The masculine reduced pronoun *ie* has a variant that is homophonous with demonstrative *die*. Since the two types of *die* are practically indistinguishable, we counted all occurrences of *die* as demonstrative.[10] In addition, when the previous word ends in /t/, *die* is sometimes also indistinguishable from *ie* (e.g., *gaat (d)ie* 'goes he'). Only cases in which a clear /d/ could be perceived were counted as *die*.

We excluded 38 cases in which participants were clearly referring either to the wrong noun phrase (which would cause a semantic mismatch in the interrogative sentence) or to something that was not mentioned in the previous sentence.[11] Non-verbatim rephrasing of the target referent was allowed (e.g., *kraan* 'crane' for *hijskraan* 'hoisting crane'). We also excluded 14 cases that contained self-repairs, and 3 cases in which the referring expression remained unclear after discussion. We excluded 4 cases in which the use of a neuter pronoun (*dat* 'that' or *het* 'it') enabled a reading in which the pronoun refers to something else than a noun phrase (e.g., to a whole proposition), and 2 cases in which participants misread (part of) the sentences such that this could

[10] We also conducted analyses on a data set in which occurrences of *die* were counted as demonstratives when stressed, and as variant of *ie* otherwise, in line with the Dutch reference grammar (Coppen et al., 2002). Stress was established by four annotators (including the author of this dissertation), who scored the degree of accentuation on each occurrence of *die* (either 'strongly accentuated', 'weakly accentuated', or 'not accentuated'). Occurrences of *die* that were agreed on by at least three of the four annotators to be unaccented were counted as 'masculine reduced pronoun'. Occurrences in which two annotators heard no accentuation but the other two did were labeled as 'unclear' and excluded from analysis. All other occurrences were categorized as 'demonstrative pronoun'. However, the distribution of unstressed *die* patterned more like stressed *die* than like reduced *ie*, being more frequent in references to objects than to subjects and not found at all in references to humans. In any case, statistical analyses showed that effects of animacy and variety were similar for the two types of categorization. Only effects of grammatical function did not reach significance anymore when occurrences of *die* were categorized differently based on stress.

[11] For pronominal references it cannot always be determined whether the participant indeed chose the correct antecedent, since masculine and feminine pronouns might be used for both the animate and the inanimate entity. These cases were therefore always treated as correct references.

influence the interpretation of the referring expression. In addition, there were 5 missing cases. In total, 66 cases (6.1%) were excluded.

## 4.5. Results

### *4.5.1. Data exploration*

Tables 4.3 and 4.4 list the frequencies of occurrence of all different types of referring expressions used by the 17 Dutch participants and by the 17 Belgian participants that scored above chance level on the grammatical gender test (henceforth 'three-way gendered Dutch'), split by the animacy and the grammatical function of the antecedent.

**Table 4.3.** Absolute and relative frequencies of referring expressions used by the NLD speakers, split by the animacy and grammatical function of the antecedent. The three pronominal categories (demonstrative, full and reduced) are broken down in the different forms within these categories (n = neuter gender; c = common gender).

| | Netherlandic Dutch | | | |
|---|---|---|---|---|
| | Animate antecedent | | Inanimate antecedent | |
| | Subject | Object | Subject | Object |
| Definite full NP | 18 (14%) | 19 (16%) | 42 (33%) | 38 (30%)[a] |
| Demonstrative | 1 (1%) | 2 (2%) | 25 (20%) | 46 (36%) |
| *dat* 'that' (n) | 0 (0%) | 0 (0%) | 1 (1%) | 1 (1%) |
| *die* 'that' (c) | 1 (1%) | 1 (1%) | 9 (7%) | 25 (20%) |
| *deze* 'this' (c) | 0 (0%) | 1 (1%) | 15 (12%) | 20 (16%) |
| Full pronoun | 85 (68%) | 83 (69%) | 41 (32%) | 28 (22%) |
| *hij* 'he' | 52 (42%) | 49 (41%) | 34 (27%) | 23 (18%) |
| *zij* 'she' | 33 (26%) | 34 (28%) | 7 (5%) | 5 (4%) |
| Reduced pronoun | 21 (17%) | 16 (13%) | 20 (16%) | 16 (13%) |
| *ie* 'he' | 1 (1%) | 0 (0%) | 8 (6%) | 6 (5%) |
| *ze* 'she' | 20 (16%) | 16 (13%) | 6 (5%) | 4 (3%) |
| *het* 'it' | 0 (0%) | 0 (0%) | 6 (5%) | 6 (5%) |
| TOTAL | 125 (100%) | 120 (100%) | 128 (100%) | 128 (100%) |

[a] Including 1 NP preceded by a demonstrative determiner.

**Table 4.4.** Absolute and relative frequencies of referring expressions used by the BD speakers split by the animacy and grammatical function of the antecedent. The three pronominal categories (demonstrative, full and reduced) are broken down in the different forms within these categories (n = neuter gender; c = common gender).

| | Three-way gendered Dutch | | | |
|---|---|---|---|---|
| | Animate antecedent | | Inanimate antecedent | |
| | Subject | Object | Subject | Object |
| Definite full NP | 1 (1%) | 0 (0%) | 4 (3%) | 2 (1%) |
| Demonstrative | 0 (0%) | 0 (0%) | 15 (12%) | 16 (12%) |
| *dat* 'that' (n) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) |
| *die* 'that' (c) | 0 (0%) | 0 (0%) | 10 (8%) | 10 (7%) |
| *deze* 'this' (c) | 0 (0%) | 0 (0%) | 5 (4%) | 6 (4%) |
| Full pronoun | 83 (65%) | 89 (68%) | 42 (32%) | 48 (36%) |
| *hij* 'he' | 56 (44%) | 58 (45%) | 21 (16%) | 24 (18%) |
| *zij* 'she' | 27 (21%) | 31 (24%) | 21 (16%) | 24 (18%) |
| Reduced pronoun | 43 (34%) | 41 (32%) | 69 (53%) | 68 (51%) |
| *ie* 'he' | 8 (6%) | 6 (5%) | 7 (5%) | 5 (4%) |
| *ze* 'she' | 35 (28%) | 35 (27%) | 61 (47%) | 62 (46%) |
| *het* 'it' | 0 (0%) | 0 (0%) | 1 (1%) | 1 (1%) |
| TOTAL | 127 (100%) | 130 (100%) | 130 (100%) | 134 (100%) |

From Tables 4.3 and 4.4 it becomes clear that both groups of speakers prefer to use full pronouns (*hij*, *zij*) to refer to animates, although the reduced feminine pronoun is also used frequently, unlike the reduced masculine pronoun. An interesting difference is that the NLD speakers use a fair amount of full NPs to refer to animates, whereas these are virtually non-existent in the BD speakers. Instead, these speakers tend to use more reduced pronouns.

When participants refer to inanimate nouns (recall that these were always feminine), the differences become more striking. The NLD speakers use all kinds of referring expressions, with a tendency away from full pronouns towards either demonstratives and full NPs or reduced (masculine) pronouns and neuter pronouns. There are even some cases in which they use a feminine pronoun. By contrast, the BD speakers show a clear preference for the reduced feminine pronoun (*ze*), which is used in almost half

of the cases. Other frequent pronouns are *zij* and, surprisingly, *hij*, the latter showing that the inanimate nouns were not always felt as feminine even by the BD speakers. Finally, there is also an increase in the use of demonstratives and full NPs for references to inanimates compared to references to animates, although not as large as for the NLD speakers.

Another striking observation is that there are only small differences between references to subjects and references to objects. The only noteworthy difference seems to be in references to inanimates, where NLD speakers use more demonstratives when referring to an object antecedent, and fewer full pronouns.

In the next three subsections, we report the effects of variety, animacy and grammatical function on the type of referring expression. We performed three logit mixed model analyses (Jaeger, 2008): one on the log odds of a personal pronoun reference (both full and reduced) out of all referring expressions (Section 4.5.2), one on the log odds of a demonstrative pronoun reference out of all pronominal expressions (Section 4.5.3), and one on the log odds of a reduced pronoun out of all personal pronouns (Section 4.5.4). animacy, grammatical function and variety were always included as fixed factors, participants and items as random factors. The fixed factors were centered to reduce collinearity between predictors. Random intercepts and random slopes for participants and items were included to account for between-participant and between-item variation.[12] Starting with a model with a full random effect structure, we used model comparisons to determine whether the inclusion of a random slope was justified by the data. Random slopes that did not contribute to the fit of the model according to a likelihood ratio test were removed (Jaeger, 2011). Only the final models are reported here.

*4.5.2. Proportion of personal pronouns out of all referring expressions*

Figure 4.1 shows the proportion of personal pronouns out of all referring expressions (personal pronouns, demonstrative pronouns and full noun phrases) as a function of the animacy and the grammatical function of the referent for NLD and BD speakers. The final logit mixed model included a by-participants random slope for grammatical function and a by-items random slope for animacy. Overall, animate referents were more likely to be pronominalized (91.8%) than inanimate referents (63.8%). This difference was significant, $\beta = 4.56$; SE = 0.71; $p < .001$. There was also a significant effect of variety, with BD speakers using more pronouns (92.7%) than NLD speakers

[12] By-participants random slopes for variety were never included, as this was a between-participants factor.

(61.9%), $\beta$ = 5.50; SE = 1.27; $p$ < .001. There was no significant interaction between animacy and variety, $\beta$ = 0.85; SE = 1.25; $p$ = .49. In addition, we found no significant main effect of grammatical function (subject: 79.2%; object: 76.0%), $\beta$ = -0.69; SE = 0.54; $p$ = .20, but there were significant interactions between grammatical function and variety, $\beta$ = -0.38; SE = 0.97; $p$ < .05, and between grammatical function and animacy, $\beta$ = -1.56; SE = 0.73; $p$ < .05.[13]

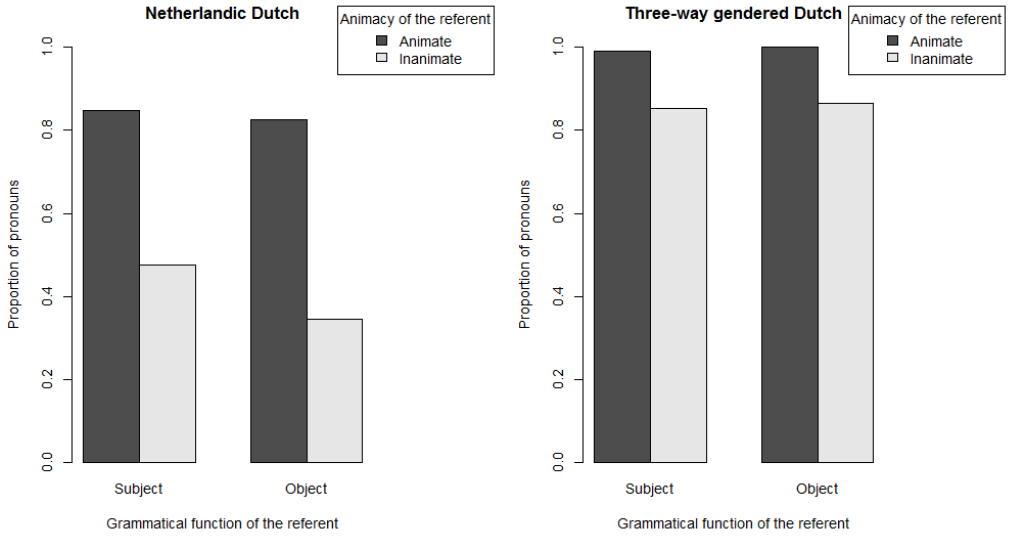


**Figure 4.1.** The proportion of personal pronouns out of all referring expressions as a function of animacy of the referent and grammatical function of the referent for NLD and BD speakers.

*4.5.3. Proportion of demonstrative pronouns out of all pronominal expressions*

Figure 4.2 shows the proportion of demonstrative pronouns out of all pronominal expressions (demonstrative pronouns and full and reduced personal pronouns) as a function of the animacy and the grammatical function of the referent for NLD and BD speakers. The final logit mixed model did not include random slopes, as they did not improve model fit. Demonstrative pronouns almost categorically referred to inanimate entities (animate referents: 0.6%; inanimate referents: 23.5%). The effect of animacy was significant, $\beta$ = -5.26; SE = 0.77; $p$ < .001. There was also a significant effect of variety, with NLD speakers using more demonstrative pronouns (19.3%) than BD speakers (6.0%). There was no main effect of grammatical function, $\beta$ = -0.47; SE = 0.78; $p$ = .55, but there was a significant interaction between grammatical function and

[13] We did not include the three-way interaction in the model due to the occurrence of empty cells.

variety, $\beta$ = -1.64; SE = 0.65; $p < .05$, suggesting that demonstrative pronouns were more likely to refer to objects than subjects, but only for the NLD speakers. The interaction between grammatical function and animacy was not significant ($p > .1$).[14]

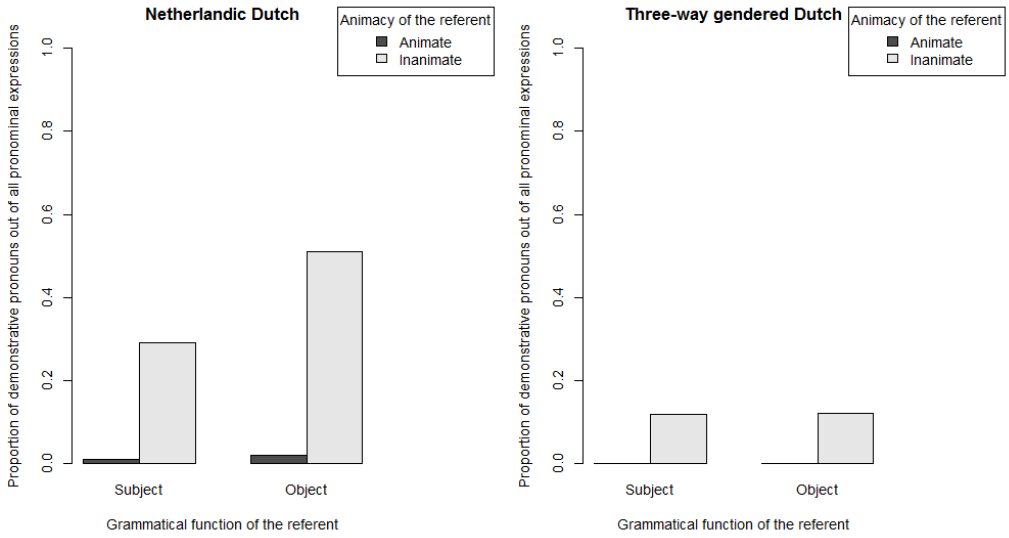


**Figure 4.2.** The proportion of demonstrative pronouns out of all pronominal expressions as a function of animacy of the referent and grammatical function of the referent for NLD and BD speakers.

*4.5.4. Proportion of reduced personal pronouns out of all personal pronouns*

We also investigated the proportion of reduced pronouns out of all pronoun references (excluding demonstratives). Feminine pronouns appeared to be more likely to be reduced (56.8%) than masculine pronouns (11.5%). As a result, any differences between the two varieties in the use of reduced pronouns could be due to the fact that BD speakers used more feminine pronouns to refer to inanimates (since all inanimate nouns in the experiment were feminine). Therefore, we analyzed masculine and feminine pronouns separately (excluding occurrences of neuter pronouns).

[14] Both the three-way interaction and the interaction between animacy and variety were not included due to the occurrence of empty cells.

Figures 4.3 and 4.4 show the proportion of masculine and feminine reduced pronoun references out of all pronoun references for animate and inanimate referents in both language varieties. Both final mixed models included no random slopes, as

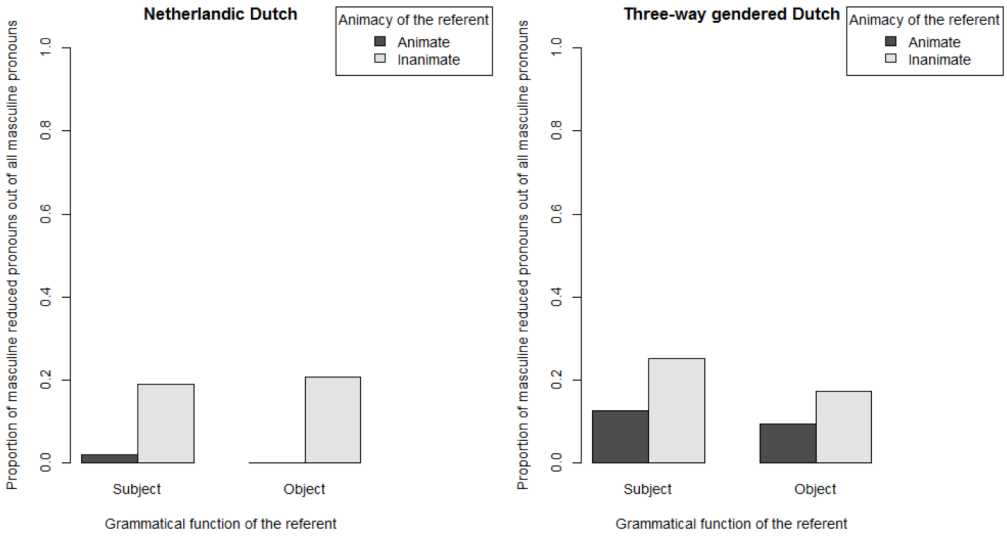


**Figure 4.3.** The proportion of masculine reduced pronouns out of all masculine pronouns as a function of the animacy of the referent and the grammatical function of the referent, for NLD and BD speakers.

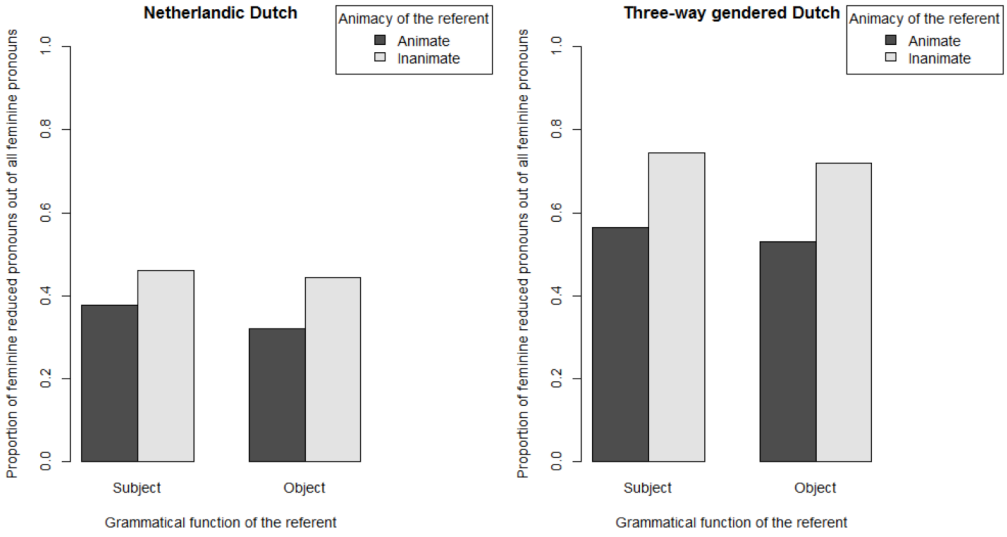


**Figure 4.4.** The proportion of feminine reduced pronouns out of all feminine pronouns as a function of the animacy of the referent and the grammatical function of the referent, for NLD and BD speakers.

these did not improve model fit. Masculine reduced pronouns were more frequent for inanimate referents (20.3%) than for animate referents (6.5%). This difference was marginally significant, $\beta$ = -3.91; SE = 2.05; $p$ = .06. There were no effects of either variety, $\beta$ = 3.07; SE = 4.69; $p$ = .51, or grammatical function, $\beta$ = 0.25; SE = 0.84; $p$ = .77 and no interactions (all $p$s > .1).[15] Feminine reduced pronouns were also more frequent for inanimate referents (70.0%) than for animate referents (45.9%). This difference was significant, $\beta$ = 1.54; SE = 0.55; $p$ < .01. Again, there were no effects of variety, $\beta$ = 2.11; SE = 1.77; $p$ = .23, or grammatical function, $\beta$ = -0.44; SE = 0.48; $p$ = .36. The interaction between animacy and grammatical function was marginally significant, $\beta$ = 1.88; SE = 1.01; $p$ = .06, suggesting that the animacy effect might be larger for object than for subject referents. The other interactions were not significant (all $p$s > .1).

## 4.6. Discussion

### *4.6.1. Likelihood of pronominalization*

In this study, we investigated the effect of animacy on the production of referring expressions in Dutch. The results show that inanimate referents are less likely to be pronominalized in Dutch than animate referents, as has been found for other languages (Dahl & Fraurud, 1996; Fukumura & Van Gompel, 2011; Yamamoto, 1999). Whereas Fukumura and Van Gompel (2011) found that speakers of English were more likely to use the third person plural pronoun *they*, which is gender neutral, for animate referents than for inanimate referents, we found the same effect for masculine and feminine singular pronouns in Dutch. This suggests that the effect of animacy is independent of whether pronouns mark gender. It also suggests that the animacy effect can be generalized to languages in which animacy is not grammatically encoded in the pronominal system, i.e., in which masculine and feminine pronouns can also refer to inanimate entities. The results further show that in references to inanimate entities the use of both demonstrative pronouns and full noun phrases increases.

The existence of two varieties of the same standard language with different gender systems allowed us to investigate possible causes of the animacy effect. According to the *conceptual accessibility account*, one possible cause is that mental representations of animate entities are more accessible in memory than representations of inanimate entities. Given that demonstrative pronouns and full noun phrases are assumed to

[15] Again, we did not include the three-way interaction due to the occurrence of empty cells.

encode a lower accessibility than personal pronouns (Ariel, 1990), our results are in line with an accessibility account.

According to the *gender avoidance account*, the fact that inanimate entities are less often pronominalized may also be due to the tendency to avoid pronouns when the gender of the antecedent noun is unclear. Since neither demonstratives nor full noun phrases mark masculine or feminine gender (demonstratives only marking common vs. neuter gender), these forms are fit for 'escaping' the gender choice that the use of personal pronouns enforces. Given that many speakers of the variety of Dutch in Belgium have clearer intuitions about masculine and feminine nominal gender than speakers of the variety of Dutch in the Netherlands, an avoidance account would predict a difference in the effect of animacy on the rate of pronominalization between Netherlandic and Belgian Dutch. Specifically, the effect of animacy should be larger for NLD speakers than for BD speakers, the latter having no reason to avoid gender-marked forms to refer to inanimates.

However, the absence of an interaction effect between variety and animacy suggests that the effect of animacy on pronominalization was equally large in the two varieties. That is, although speakers of Belgian Dutch pronominalized referents more frequently than speakers of Netherlandic Dutch, inanimate referents were less likely to be pronominalized than animate referents to the same degree as in Netherlandic Dutch. Hence, the animacy effect cannot just be due to an avoidance strategy in the NLD speakers.

Because we would expect variation in the use of referring expressions given the linguistic contexts as used in the present experiment, the finding that the BD speakers used a lot of personal pronouns (and few demonstratives and full noun phrases) in comparison to the NLD speakers is unexpected. This might be due to participants' idiosyncratic preferences or task strategies. For instance, of the BD speakers, 10 out of 17 always used pronouns in the experimental items. Conversely, two NLD speakers always used pronouns, while one participant always used full NPs. It might be the case that once a type of referring expression was chosen, these participants stuck to it throughout the experiment. However, removing the participants that showed no variation in the choice of referring expression did not change the pattern of results.

It might also be the case that the effect is related to a greater sensitivity of Belgian Dutch speakers to the grammatical gender of the nouns. That is, since there is a one-to-one mapping between nominal and pronominal gender for these speakers, pronouns may arise more naturally than in Netherlandic Dutch, even in animate

reference. This explanation seems unlikely, however, since there are no other indications that Dutch speakers from Belgium would generally use pronouns more often than Dutch speakers from the Netherlands. (e.g., a quick search in the Corpus of Spoken Dutch (*Corpus Gesproken Nederlands*; CGN) did not reveal any such difference). Alternatively, the more frequent use of full NPs by the NLD speakers might be regarded as a gender avoidance strategy after all. However, it is not clear why such a strategy would also occur with animate antecedents, as the natural gender of these antecedents was always clear, and thus there is no reason why pronouns should have to be avoided.

Finally, the difference between the two groups might be due to the fact that the BD speakers were somewhat older on average than the NLD speakers. Perhaps the likelihood of pronoun use increases with age (see, e.g., Hendriks et al., 2013). However, a logit mixed model with age as predictor did not reveal a significant effect on pronoun use. More data may be needed to resolve this issue.

Another surprising finding is that there was no main effect of grammatical function of the antecedent on pronoun use. Grammatical function is generally assumed to be an important factor in determining the accessibility of the referent (notably within *centering theory*, Gordon et al., 1993; Grosz et al., 1995). Fukumura and Van Gompel (2011) also found that the subject of the previous sentence was more likely to be pronominalized than the object. The reason why we did not find such an effect may be related to task-specific factors. For example, our participants had to speak aloud, while other studies have often used written sentence completion tasks.

Still, there did seem to be an effect of grammatical function within references to inanimates in the present study, at least for the Netherlandic Dutch speakers (see Table 4.3 above). Here, references to objects showed a decrease in the use of full pronouns (notably *hij* 'he') with respect to references to subjects, and an increase in the use of demonstratives (notably *die* 'that'). This is in line with findings that *hij* has a preference to refer to salient entities, while *die* refers to less salient entities (e.g., Kaiser & Trueswell, 2004). The fact that these differences only occurred for references to inanimates suggests that animacy overruled the effect of grammatical function: Human entities are preferred to be pronominalized, regardless of whether they are mentioned in subject or in object position. The absence of the effect of grammatical function in (three-way gendered) Belgian Dutch might be due to the strong preference of these speakers to use personal pronouns, which could have masked effects of grammatical function.

*4.6.2. Use of full and reduced pronouns*

The results of the present study further show that when pronominalized, inanimate nouns are more likely to be referred to with the reduced forms than animate nouns. While this finding is in line with the claim that full pronouns typically refer to (human) animates (Cardinaletti & Starke, 1996), this seems more a statistical tendency than a categorical rule, given that full pronouns were also used to refer to inanimate entities in the experiment. Conversely, while the full pronoun *hij* 'he' is classified as an exception in the Dutch reference grammar in that it can refer to inanimate entities just as well as to animate entities (Coppen et al., 2002: §5.2.5.2.1), our results show that also in this case the full pronoun is more frequent in reference to animates than in reference to inanimates.[16] In addition, we found a few occurrences of the neuter pronoun *het*, which does not have a full form (i.e., it cannot be stressed, Coppen et al., 2002). All of these occurrences were references to inanimate entities.

The finding that reduced pronouns are used more often to refer to inanimate entities seems to be incompatible with the conceptual accessibility account. This account states that more attenuated forms typically mark entities that are more accessible in memory (e.g., Ariel, 1990; Gundel et al., 1993). In practice, this means that speakers will be more likely to choose more attenuated expressions for more accessible referents. For example, when a referent is given or predictable from the context, its mental representation is already highly activated, and the speaker may suffice with an attenuated referring expression. When the referent is new or unpredictable, the speaker marks the fact that there is not yet an activated memory representation by using a more elaborate expression. Since animate entities are assumed to be more accessible than inanimate entities, an accessibility account predicts that reduced (i.e., more attenuated) pronouns should be more likely to refer to animate entities. This is the opposite of what we found. Our finding is in line with Kaiser (2011) and Kaiser and Trueswell (2004), who found that the distribution of full and reduced pronouns in Dutch cannot be explained by an accessibility account.

We offer two possible (not mutually exclusive) explanations for the effect of animacy on the use of reduced pronouns in Dutch: Firstly, it could be the case that the full forms of the third person personal pronoun emphasize the encoded gender information more than the reduced forms. Thus, by using *hij* or *zij* a speaker stresses

[16] Note that this probably only holds in inversion contexts such as in the present experiment, since the reduced pronoun *ie* cannot occur in sentence-initial position. In that position, the full form has to be used.

that the intended antecedent should be a masculine or a feminine entity, respectively. Especially with the loss of grammatical gender in the nouns (but perhaps also for speakers that retain it), these pronouns may be more likely to be associated with natural gender. This makes them less suitable for reference to inanimates, which do not have a natural gender. The use of a reduced form might mitigate this association, and hence may be better suited for inanimate referents.

Secondly, it could be the case that full pronouns convey more important information than reduced pronouns. Often, importance coincides with accessibility: What is accessible is typically less important, because it is already predictable. The two notions are not necessarily two sides of the same coin, however. For example, Watson, Arnold, and Tanenhaus (2008) found that important information was pronounced with a higher intensity than less important information, independently of predictability. In the same vein, important information may be uttered with more prominent forms, regardless of its accessibility status.

Indeed, there is evidence that full pronouns in Dutch are used to refer to more important information. For example, in a corpus study, Bouma (2008) found that full pronouns were more frequent than reduced pronouns in the preverbal position in Dutch main clauses, which he argues is a position in which important information in expressed. In addition, in a story completion experiment conducted in Dutch, Vogels, Krahmer, and Maes (2011) showed a tendency for full pronouns to be more frequent with visually salient referents, which could also be associated with high informational importance. Finally, Kaiser and Trueswell (2004) proposed, based on eye-tracking and corpus data, that contrast might be a critical factor in the distribution of full and reduced pronouns in Dutch: While reduced pronouns preferably refer to the subject of the previous sentence, full pronouns are used to signify a contrast between two or more possible referents. This use is independent of accessibility: Even though the antecedent of the pronoun may be highly accessible (which should be a prerequisite for using a pronoun in the first place), the full form rather than the reduced form is chosen to emphasize the contrasting information, i.e., the information that the speaker would like to bring forward as important (e.g., Chafe, 1976).

Kaiser and Trueswell (2004; see also Kaiser, 2010, 2011; Kaiser & Trueswell, 2008) suggest that it is probably a combination of multiple factors that determines the use of full and reduced pronouns, contrast being one of them. The present study suggests that animacy may be another one. Animate entities may be inherently more important than inanimate entities (e.g., Givón, 1983; Kirsner, 1979), for example because they are

typically the instigators of events, or because we empathize with them (e.g., Kuno & Kaburaki, 1977). Therefore, pronoun references to animate entities may be less reduced than pronoun references to inanimate entities.

The apparent asymmetry between the choice to pronominalize (animate referents are more often pronominalized) and the choice to produce a reduced pronominal form (pronouns referring to animate referents are less often reduced) is in line with models of reference production that allow multiple factors to have different effects on specific referential choices. Based on the distribution of pronouns and demonstratives in Finnish, Kaiser and Trueswell (2008) argue for a *form-specific multiple constraints approach* of reference resolution, in which the interpretation of one form may be more sensitive to certain factors (e.g., grammatical function, information structure) than the interpretation of another form. On the production side, the choice to pronominalize may be most sensitive to the referent's accessibility or salience, while the choice between full and reduced forms is more sensitive to importance or newsworthiness of information, at least in Dutch.

The choice between a full and a reduced pronoun may be related to the degree of acoustic reduction in reference production. In line with findings that words and syllables become more reduced when they are more predictable based on the linguistic context (e.g., Aylett & Turk, 2004; Bell, Jurafsky, Fosler-Lussier, Girand, Gregory, & Gildea, 2003) or based on the nature of the task (Watson et al., 2008), Kaiser, Li, and Holsinger (2011) found evidence that names referring to more predictable referents (referents that had previously been found to have a higher likelihood of subsequent mention) were shorter than names referring to less predictable referents. Crucially, in Kaiser et al.'s experiments predictability did not influence the choice between a name and a pronoun. Although the distinction between full and reduced pronouns in Dutch is not merely a matter of acoustic reduction (they are separate lexical forms), these findings do support our conjecture that the choice between a pronoun and a more specific expression and the choice between a full and a reduced pronoun are separate referential choices that are driven by different factors.

*4.6.3. Open issues*

A number of questions remain. First of all, the loss of the distinction between masculine and feminine nouns in Netherlandic Dutch and its preservation in Belgian Dutch possibly also has consequences for pronoun ambiguity. For example, in

contexts with both an animate and an inanimate possible referent, a pronoun is ambiguous when the referents have the same grammatical gender. For BD speakers, this may lead to the avoidance of pronouns in favor of more specific referring expressions, since the same pronoun can easily refer to either the animate or the inanimate entity. For NLD speakers, a pronoun should typically refer to the animate entity in same-gender contexts, regardless of whether it could technically also refer to the inanimate entity.

To investigate these predictions, we tested the effect of same-gender contexts on pronoun avoidance in the present data set. Since in our experimental materials pronouns were always eventually disambiguated, we only investigated filler items in which pronouns remained ambiguous. Only in items that contained two animate entities, and only in the Netherlandic Dutch data, we found fewer pronouns in same-gender contexts (73.8%) than in different-gender contexts (83.1%), and more demonstratives (6.2% vs. 3.4%) and full NPs (20.0% vs. 13.6%). Although this difference was not present in the Belgian Dutch data, it suggests that people are more sensitive to ambiguity when two possible referents have the same *natural* gender. Another possibility is that referents of the same animacy are conceptually more similar, which decreases the accessibility of both referents, because they compete more for attention (Arnold & Griffin, 2007; Fukumura & Van Gompel, 2011). Further research is needed to determine whether and how similarity in grammatical gender affects accessibility and pronoun use.

Secondly, it is an open question whether the effects that we found are speaker-internal or arise from addressee-oriented processes. The avoidance of pronouns due to the speaker's uncertainty about grammatical gender seems to be a clear case of a speaker-internal process. Conceptual accessibility effects on language production have also generally been described as emerging from the speaker's own memory (e.g., Bock et al., 1992; Fukumura & Van Gompel, 2011; Montag & MacDonald, 2013; Prat-Sala & Branigan, 2000). For example, mental representations of animate entities are more easily retrieved, which in turn affects the language production process. In the case of referring expressions, Fukumura and Van Gompel (2011) argue that animate entities need less encoding of conceptual information because of their higher conceptual accessibility. Thus, speakers use shorter referring expressions when referring to animates than when referring to inanimates.

Alternatively, it could also be the case that speakers are aware that addressees may be focusing more on animate than on inanimate entities, for example because animate

entities are more likely topics. Hence, addressees expect speakers to be more likely to talk about animate entities, and speakers accommodate this expectation by using more specific expressions when it is not fulfilled (i.e., when they talk about an inanimate entity; cf. Arnold, 2010). In this case, the locus of the animacy effect is in the speaker's beliefs of the referent's accessibility in the addressee's memory. More research is needed to separate these alternative explanations.

Thirdly, a potential complication of our results is that we found that even our BD speakers sometimes used the 'wrong' gender. For example, they still used a masculine pronoun (*hij* or *ie*) to refer to inanimate entities (which should all have been grammatically feminine) in over 20% of the cases. In addition, the results of the post-hoc gender test showed that not all BD speakers employed a nominal three-way gender distinction. Hence, the question arises whether the BD speakers were not also avoiding gender-marked expressions for inanimate antecedents, which might explain why we did not find differences in the effect of animacy between the two varieties.

The first thing to be noted here is that in the post-hoc gender test, especially participants below the age of 50 used *hij* to refer to feminine nouns. This is in line with findings that the three-way gender system in Belgian Dutch is in the process of erosion as well, and that children who are learning the language often choose a pronominal form based on semantic rather than grammatical grounds (De Vogelaer, 2006; De Vogelaer & De Sutter, 2011). The older participants in the test did make use of the nominal gender distinctions, although none of them performed perfectly according to the original gender of the nouns (as found in the dictionary and our pretest). However, this gender task was a conservative test, since all items had the same expected answer (i.e., feminine pronouns). Therefore, participants could have been reluctant to choose the same answer on each trial, adding some variation, unless they were really confident about their answer.

In addition, our BD speakers were not a homogeneous group with respect to their dialect backgrounds. While all southern dialects still make a three-way distinction in the nominal genders, they may differ in whether a noun is assigned masculine, feminine or neuter gender (De Vogelaer, 2009). Thus, although all inanimate nouns were pretested for grammatical gender by BD speakers (also from different backgrounds), we cannot exclude the possibility that the grammatical gender of the words used in the experiment differed across participants due to their different backgrounds. Apart from these considerations, although there may be an increasing uncertainty about nominal gender also among speakers of Belgian Dutch, the

difference with Netherlandic Dutch can be assumed to be still significant, with especially the older generations keeping to the three-way gender system. For example, De Vogelaer and De Sutter (2011) report a rate of 83.6% use of grammatical agreement for Flemish speakers above the age of 55. Hence, we assume that our BD speakers were at least more certain about grammatical gender than our NLD speakers.

Fourthly, it could be the case that gender avoidance is more common in written language than in spoken language: Especially in more formal texts, writers might be more stimulated to 'get the gender right', and hence will avoid pronouns when they are not sure about the gender of a noun (Audring, 2009). This might be less of an issue in spontaneous spoken language. Note, however, that while we elicited spoken responses in our experiment, they were not spontaneous. In fact, because participants were in a test-like situation, in which they had to fill in gaps in sentences, they might have been especially encouraged to 'get the gender right'. Hence, we should have expected a lot of gender avoidance in references to inanimate antecedents.

On the other hand, the attempt to choose the correct gender might explain why some Dutch participants also sometimes used feminine pronouns to refer to inanimate entities (although not systematically), whereas this would not be expected in spontaneous language production. This finding might also illustrate the complete helplessness of the Dutch participants when it comes to grammatical gender, switching back and forth between all kinds of referring expressions. Furthermore, the use of full and reduced pronouns is also likely to be affected by register: The reduced pronoun *ie* is highly colloquial, whereas *ze* is not, although in formal written text *zij* may be more frequent (e.g., Van Bergen et al., 2011). Given these issues, at least the difference between the spoken and the written modality should be researched further.

Finally, an interesting remaining question is whether grammatical gender also directly affects the conceptual accessibility of referents. It has been found that people conceptualize objects differently based on the grammatical gender of the corresponding noun (see Boroditsky, Schmidt, & Phillips, 2003 for a review). For example, the word for 'bridge' is masculine in Spanish and feminine in German. When describing bridges, German speakers were found to use more adjectives denoting typical feminine properties, such as 'beautiful' and 'elegant', while Spanish speakers used more words such as 'dangerous', and 'strong', which were rated as typically masculine (Boroditsky, Schmidt, & Phillips, 2002, as cited in Boroditsky et al., 2003). This suggests that referents of feminine nouns are conceptualized as more feminine, and referents of masculine nouns as more masculine. More generally, then,

it might be the case that using 'animate' genders (masculine or feminine; Dahl, 2000) for inanimate nouns causes these inanimate entities to be conceptualized as more animate. This could be a topic for future research.

### 4.7. Conclusions

The study presented in this chapter investigated effects of animacy on pronoun use in Belgian and Netherlandic Dutch. Although speakers of Belgian Dutch appeared to use pronouns more frequently in general, animate entities were more likely to be pronominalized than inanimate entities, as was found for speakers of Netherlandic Dutch. We conclude that the tendency in Dutch to avoid the use of pronouns when referring to inanimate entities cannot solely be due to a strategy to avoid a gender choice, and is compatible with the general preference to use less attenuated forms for less conceptually accessible entities. This tendency is however countered by a preference for full over reduced pronouns when referring to animate entities, which may be related to the importance or newsworthiness of the referent rather than to accessibility.

# Chapter 5
# Cognitive load

**Abstract**

We report on two experiments investigating the effect of an increased cognitive load for speakers on the choice of referring expressions. Speakers produced story continuations to addressees, in which they referred to characters that were either salient or non-salient in the discourse. In Experiment 1, referents that were salient for the speaker were non-salient for the addressee, and vice versa. This setup was compared to a situation in which all discourse information was shared in Experiment 2. Cognitive load was manipulated by the presence or absence of a secondary task for the speaker. The results show that speakers under load are more likely to produce pronouns, at least when referring to less salient referents. We take this finding as evidence that speakers under load have more difficulties taking discourse salience into account, resulting in the use of expressions that are more economical for themselves.

**This chapter is based on:**

Vogels, J., Krahmer, E. J., & Maes, A. A. (submitted). How cognitive load influences speakers' choice of referring expressions. Submitted for journal publication.

## 5.1. Introduction

When speakers refer to something, they have to choose a certain type of referring expression, such as a definite description (e.g., *the girl*) or a pronoun (e.g., *she*). Traditionally, the speaker's choice of a referring expression has been assumed to be tailored for the addressee (e.g., Ariel, 1990; Gundel, Hedberg, & Zacharski, 1993). According to this view, speakers make assumptions about the cognitive status of the referent in the mind of their addressee. The most important factor in determining this status is assumed to be the salience of the referent in the discourse. For example, if the referent was the topic of the previous sentence, it can be assumed to be highly accessible in the addressee's discourse model (e.g., Givón, 1983). Therefore, it does not need an elaborate description to be reactivated in the memory of the addressee. Because, so the classical reasoning goes (e.g., Gundel et al., 1993), the addressee knows that the speaker would have used a more elaborate expression if she had a less activated referent in mind, the use of an attenuated expression, such as a pronoun, aids the addressee's interpretation. This is in line with the idea that cooperative speakers obey Grice's Maxim of Quantity (Grice, 1975), i.e., speakers choose referring expressions that are as informative as required for the addressee to pick out the correct referent, but not more informative than required.

More recently, it has been suggested that the choice of a referring expression may also be influenced by speaker-internal constraints (e.g., Arnold, 2008; Arnold, Bennetto, & Diehl, 2009; Arnold & Griffin, 2007). Speakers are not always monitoring the communicative needs of their addressees (e.g., Dell & Brown, 1991; Engelhardt, Bailey, & Ferreira, 2006; Wardlow Lane, Groisman, & Ferreira, 2006). One reason is that the language production system is constrained by the speaker's attention resources and working memory capacity. Since these resources are limited (Baddeley, 1986), speakers do not have unrestricted processing capacity to keep track of all elements in the discourse and to calculate the accessibility of referents for the addressee. In addition, it has become clear that even when there is sufficient processing capacity, people do not always initially take the perspective of their conversation partners into account in producing referring expressions (e.g., Bard, Anderson, Sotillo, Aylett, Doherty-Sneddon, & Newlands, 2000; Fukumura & Van Gompel, 2012; Gann & Barr, 2012), although they might do this eventually, in a later stage of processing (e.g., Dell & Brown, 1991; Horton & Keysar, 1996; Keysar, Barr, & Horton, 1998).

It is less clear, however, how exactly speaker-internal constraints affect reference production. In this chapter, we experimentally investigate how an increased memory and attention load on the part of the speaker influences how speakers choose between attenuated expressions such as pronouns, and more elaborate expressions such as full noun phrases. Assuming that such cognitive load taps into the language production process, manipulating it gives more insight in the mechanisms underlying referential choices. The next sections discuss different hypotheses with respect to the possible effects of cognitive load on the choice of referring expression. On the one hand, if speakers are choosing referring expressions based on their assumptions about the referent's accessibility in the addressee's model of the discourse, increased cognitive load may cause speakers to be less able to make these assumptions (Section 5.2). On the other hand, if speakers are choosing referring expressions based on how much attention they themselves allocate to the referent, increased cognitive load may make speakers less attentive, and hence may affect the referent's representation in their own discourse model (Section 5.3).

## 5.2. Hypothesis 1: Cognitive load makes reference more egocentric

If choosing referring expressions involves taking into account how accessible the referent is for the addressee, an increased cognitive load may make this harder. Different models of audience design have been proposed to account for the fact that speakers are not always monitoring the knowledge of their addressee. According to the Monitoring and Adjustment model (Horton & Keysar, 1996), for example, speakers initially plan their utterances egocentrically, i.e., without taking into account common ground with their addressee. A subsequent process then checks this initial plan for errors, such as whether it is relying on information that is not accessible for the addressee, and adjusts it when necessary. Since this monitoring involves an additional step in processing, it is predicted to take up more time and memory resources. Indeed, Horton and Keysar (1996) found that speakers took into account the addressee's perspective when they had to choose whether or not to include an adjective in their referring expressions. However, they were less able to do this when they were under time pressure. In that case, speakers more often based their utterances only on information accessible to them.

In addition, the Dual Process model (Bard et al., 2000; Bard & Aylett, 2005) proposes that a distinction should be made between automatic processes that only take into account the speaker's knowledge, and more effortful processes that build inferences about the addressee's knowledge. These inferential processes compete for attention with task demands: The more attention a task requires, the less speakers will take the addressee's knowledge into account.

Thus, both the Monitoring and Adjustment model and the Dual Process model suggest that restrictions on the processing capacity needed for audience design may make references more egocentric. However, it is not completely clear what it means to be egocentric when choosing between pronouns and full noun phrases. On the one hand, it could mean that speakers base the choice of referring expressions on their own model of the discourse rather than on assumptions about the addressee's discourse model (e.g., Bard & Aylett, 2005; Fukumura & Van Gompel, 2012). On the other hand, referring egocentrically could mean that speakers are inclined use those referring expressions that are easiest for them to produce (e.g., pronouns; Hendriks, Englert, Wubs, & Hoeks, 2008). If the choice of referring expression becomes more based on the speaker's own model of the discourse when memory load is high, speakers are expected to be more likely to use a pronoun when the referent is highly accessible for them, and a full noun phrase when the referent is less accessible for them, rather than consider accessibility from the addressee's perspective. In many cases, of course, this will not cause problems, since speakers and addressees tend to have closely aligned discourse models (e.g., Arnold, 2008; Pickering & Garrod, 2004); a referent that is highly accessible in the speaker's discourse model is typically also highly accessible in the addressee's discourse model. However, when the speaker's and the addressee's perspectives differ – for instance, because the addressee did not hear part of the preceding discourse – speakers under load might be inclined to use pronouns if the referent is salient in their own discourse model but not necessarily salient in the addressee's discourse model. Conversely, they might be inclined to use full noun phrases if the referent is not salient in their own discourse model but salient in the addressee's discourse model.

To investigate whether speakers are taking into account their addressee's perspective when choosing referring expressions, Fukumura and Van Gompel (2012) conducted a story completion experiment in which the sentence directly preceding the speaker's continuation was in privileged ground, i.e., it was only heard by the speaker over headphones. This privileged sentence either made the referent to be described in

the continuation (the target referent) discourse salient, or it made a competitor referent salient. In both cases, the target referent was not salient for the addressee. Therefore, if speakers were taking into account the addressee's discourse model, they should use a full noun phrase to refer to the target referent, irrespective of the content of the privileged context sentence. If speakers were using their own discourse model, they should use more pronouns when the target referent was made salient in the privileged sentence than when it was not. The results of this study showed the latter pattern, suggesting that speakers were more likely to follow their own discourse model than to take into account their addressee's perspective. However, these results were not completely identical to the findings from another condition in which all discourse context was in common ground, which might suggest that some audience design was going on. One possibility is that under load, speakers are more likely to abandon such audience design, and choose referring expressions based on their own discourse model.

Referring egocentrically may also mean that speakers choose more economical expressions overall. Due to their reduced phonological and semantic content, pronouns may be more economical in terms of processing costs than full noun phrases (Almor, 1999; Burzio, 1998; Levinson, 1987). Therefore, speakers may inherently prefer to produce pronouns over more specific expressions. On this view, given that speakers themselves know what they are referring to, producing more elaborate expressions is simply not beneficial for speakers. Any expression that is less economical than a pronoun may thus be considered as somehow tailored for a (potentially hypothetical) addressee (Hendriks, Koster, & Hoeks, 2013). If this addressee-oriented process is cognitively effortful, the preference to use pronouns may be reinforced when speakers do not have enough processing capacity to take into account the knowledge of the addressee. Indeed, studies have found that speakers with a low working memory capacity (children, elderly) are more likely to use pronouns in contexts in which the referent is not salient for the addressee (and hence a more specific expression would normally have been appropriate; Hendriks et al., 2008; 2013; Wubs, Hendriks, Hoeks, & Koster, 2009; see also Almor, Kempler, MacDonald, Andersen, & Tyler, 1999).

In sum, our first hypothesis is that that cognitive load will make speakers more egocentric. This could result in either choosing referring expressions based on the speaker's own discourse model (i.e., producing pronouns for referents salient for the

speaker and full noun phrases for non-salient referents), or generally using more economical expressions (i.e., pronouns).

### 5.3. Hypothesis 2: Cognitive load affects the speaker's own discourse model

If speakers are using their own discourse model when choosing referring expressions, rather than taking into account the discourse model of the addressee, regardless of whether they have enough processing capacity for that, cognitive load may directly affect the accessibility of mental representations in the speaker's own memory (Arnold, 2010). Indications that the activation of discourse elements in the speaker's own discourse model influences the choice of referring expression come from studies that manipulate the speaker's attention resources. For example, Arnold and Griffin (2007; see also Fukumura, Van Gompel, & Pickering, 2010) conducted a story completion experiment in which they varied the number of possible referents in the discourse. They found that speakers used fewer pronouns when a referential competitor was present, even though pronouns were never ambiguous and the target referent was salient in the discourse. Hence, a pronoun reference could have been easily resolved by the addressee. Speakers have also been found to use fewer attenuated expressions when they are distracted by another task (Rosa & Arnold, 2011), and when they are either disfluent or planning longer utterances, which are both considered indications for an increased cognitive load (Arnold et al., 2009). These findings have been explained as evidence for a decrease in the accessibility of the referent in the speaker's own discourse model when attentional resources have to be spread over multiple possible referents or multiple (effortful) tasks. Thus, on this view, speakers with decreased cognitive resources are less likely to use attenuated expressions, because the activation of the referents in their own discourse model is reduced.

Alternatively, cognitive load may affect the degree to which speakers can keep track of the salience of referents in their own discourse model. In this case, an increased cognitive load may result in less consistent use of referring expressions (Arnold, 2010; Hendriks et al., 2013). That is, pronoun use may become less tied to the discourse context. For example, speakers under load may use fewer pronouns for referents that are salient in the discourse, and more pronouns for non-salient referents, compared to the normally predicted consistent use of pronouns for salient entities and full noun

phrases for non-salient entities. Indeed, Hendriks et al. (2013) found that although elderly adults were capable of taking into account the accessibility of the referent for the addressee, they still used more pronouns for non-salient referents than younger adults. This suggests that they had difficulties keeping track of the salience of the discourse referents.

In sum, our second hypothesis is that speakers base their choice of referring expressions on the salience of the referent in their own model of the discourse. In this case, an increased cognitive load may reduce referent accessibility, resulting in more elaborate expressions, or it may make it more difficult for the speaker to keep track of the discourse, resulting in less consistent use of referring expressions.

### 5.4. Predictions and experimental design

We have formulated two, not mutually exclusive, hypotheses with respect to the effect of an increased cognitive load on the speaker's choice of referring expressions. Firstly, when they are under load, speakers may be less likely to take into account the addressee's perspective, causing them either to fall back on their own discourse model or to use more economical expressions. Secondly, cognitive load may affect mental representations in the speaker's own discourse model, either resulting in a reduced accessibility of the referents, and hence in more elaborate referring expressions, or leading to less consistent use of referring expressions.

To tease these possible effects of cognitive load apart, the speaker's and the addressee's perspectives with respect to the discourse salience of the referent should be dissociated. In this way, it can be determined whether speakers under load have more difficulties to choose referring expressions based on their addressee's perspective. In addition, references to both salient and non-salient referents should be investigated, because cognitive load might affect these differently. For example, whereas a reduced accessibility is predicted to decrease pronoun use for both salient and non-salient referents, having difficulties in keeping track of the discourse model predicts that pronoun use decreases for salient referents, but increases for non-salient referents (i.e., choice of referring expression becomes less consistent). Indeed, most studies that found a decrease in pronoun use with increased cognitive load (e.g., Arnold & Griffin, 2007; Fukumura et al., 2010; Rosa & Arnold, 2011) investigated only contexts in which the target referent was the subject of the preceding sentence, and

was therefore salient for the speaker (and the addressee, whether hypothetical or not). By contrast, studies that found an increase in pronoun use in speakers with a lower working memory capacity either investigated only non-salient referents (i.e., after a topic shift; Hendriks et al., 2008; Wubs et al., 2009) or did not control for discourse salience (Almor et al., 1999). None of these studies dissociated the speaker's and addressee's perspectives, by which it remains unclear whether cognitive load was affecting referent accessibility in the speaker's own discourse model or her assumptions about the accessibility of the referent in the addressee's discourse model. The present study was conducted to shed more light on this question.

We conducted two story completion experiments in Dutch, with a similar setup as in Fukumura and Van Gompel (2012). We manipulated the cognitive load of the speaker, as well as referent salience and perspective. In the experiments, speakers were presented with pairs of pictures showing two characters. Following two context sentences, speakers orally produced a continuation for an addressee, in which they referred to one of these characters, which was either salient or non-salient in the discourse. We manipulated cognitive load by having speakers conduct a verbal memory task while telling the stories in one half of the experiment. We used a verbal rather than a visual secondary task (cf. Rosa & Arnold, 2011) to make sure that it would interfere with memorizing or attending to discourse information rather than with visually attending to the characters in the pictures (e.g., Kellogg, Olive, & Piolat, 2007). In Experiment 1, we dissociated the perspectives of the speaker and the addressee, such that whenever the referent was salient for the speaker, it was non-salient for the addressee, and vice versa. Perspective was manipulated by presenting the context sentence directly preceding the speaker's continuation only to the speaker, over headphones, as in Fukumura and Van Gompel (2012). To be able to compare this setup with a situation in which all discourse information was shared between speaker and addressee, Experiment 2 tested the effect of cognitive load without a perspective difference.

## 5.5. Experiment 1

### *5.5.1. Methods*

#### 5.5.1.1. Participants

Sixty-four students (47 female; mean age 20.2 years) from Tilburg University participated in the experiment for course credit. Half of them acted as speakers, the others acted as addressees. All were native speakers of Dutch.

#### 5.5.1.2. Materials

The experimental items consisted of 16 pairs of photographs, taken from the materials used in Chapter 2, accompanied by two introductory sentences and the onset of a third sentence. The first picture of a pair always showed one male and one female person sitting next to each other. In the second picture, one of these persons performed an action, such as walking away or getting a glass of water. This person will be referred to as the target character, as participants were expected to refer to this character in their continuations. There were two versions of each picture pair; one in which it was the male person and one in which it was the female person that performed the action. An example of a picture pair is shown in Figure 5.1.[1]

The first sentence introduced both characters with indefinite noun phrases, which were either *een meisje* 'a girl' and *een jongen* 'a boy', or *een vrouw* 'a woman' and *een man* 'a man'. One of these was mentioned as the subject, and the other in a prepositional phrase (e.g., *Een meisje zat te discussiëren met een jongen* 'A girl was arguing with a boy'). This sentence was read aloud by the speaker to the addressee. The second sentence described some emotional or physical state of the person mentioned in the prepositional phrase (e.g., *De jongen raakte enorm gepikeerd* 'The boy got really annoyed'). Hence, there was always a topic shift between the first and the second context sentence. The second sentence was prerecorded by a female native speaker of Dutch and was only heard by the speaker over headphones. The onset of the third sentence was always *Vervolgens...* 'Subsequently...', serving as a cue for the speaker to complete the story based on the second picture. In the speaker-salient

[1] As discussed in Chapter 2, these stimuli were originally developed to investigate the influence of visual salience on reference production, hence the visual fore- or backgrounding of one of the characters in Figure 5.1. In the present experiment, visual salience was counterbalanced across items, i.e., in half of the items the target referent was in the foreground in condition A and in the background in condition B, and in the other half it was the other way around.

condition (condition A in Figure 5.1), the target character in the second picture was the subject of the privileged sentence, and therefore discourse salient for the speaker but not for the addressee. In the addressee-salient condition (condition B in Figure 5.1), the target character was the subject of the introductory sentence, in which case it was discourse non-salient for the speaker but salient for the addressee, since this sentence was the only context sentence heard by the addressee.


**Figure 5.1.** Example of a stimulus item in two conditions in Experiment 1. Sentence 1 was read aloud by the speaker; sentence 2 was presented only to the speaker, over headphones. Context sentences are translations of the Dutch originals.

In addition, 20 picture pairs served as fillers. These differed from the experimental items in that some showed either two male or two female characters (and hence pronouns would be ambiguous) or only one character. In this way, using only one type of expression throughout the experiment was discouraged. In the accompanying sentences, some characters were given labels such as *een verkoopster* 'a saleswoman' or *een Duitser* 'a German', and sometimes the same character was the subject of both introductory sentences. An additional 4 items were included as practice items.

5.5.1.3. Procedure

The experiment took place in a quiet room. Two participants were randomly assigned to the role of speaker and addressee. The participant taking the role of speaker was seated at one end of a table, behind a laptop connected to a PST Serial Response Box. The participant taking the role of addressee was seated at the other end of the table, and was given a booklet containing all different picture pairs and an answer sheet.

The experiment was run on the laptop using E-Prime 2.0 software, and was only visible to the speaker. The speaker's task was to complete the stories depicted by the picture pairs in such a way that the addressee could pick out the correct picture pair from the booklet.

Crucially, in half of the trials, the speaker received a secondary task (cognitive load condition), while there was no secondary task in the other half (no cognitive load condition). In the no cognitive load condition, each trial started with the item number presented on the screen, accompanied by a 500 ms beep, followed by a fixation cross. Then, the first picture of a pair appeared on the left side of the screen. After 3 s, the first introductory sentence appeared below the picture in a red font. The speaker read this sentence aloud to the addressee. After 5 s, the second sentence was presented to the speaker over the headphones. Next, while the first picture remained visible, the second picture appeared automatically on the right side of the screen, together with the onset of the third sentence, which also appeared below the picture in a red font. At this time, recording started, and the speaker had 6 s to complete the story based on the event shown in the picture, by saying it aloud to the addressee. When this interval had elapsed, recording stopped and the pictures and sentences disappeared. The addressee's task was to select the correct picture pair out of three options from the booklet and mark the correct answer on the sheet. In the experimental items, two of the three options differed only in which character performed the action, making correct reference crucial for the addressee to finish his task successfully. The addressee was instructed to give the speaker a hint when the next trial could be started.

In the cognitive load condition, the appearance of the first picture was preceded by the words BAL or DAL (Dutch for 'ball' and 'valley', respectively)[2], which was presented in the middle of the screen for 1 s. The same happened at the end of the trial, followed by the question *Was dit woord hetzelfde als het vorige woord? (Ja/Nee)* 'Was this word the same as the previous word? (Yes/No)'. The speaker then pressed either the green/Yes or the red/No button on the response box. They did not receive feedback on their answers.

The participants received instructions both orally and in written form. Speakers were explicitly told that the sentence presented over headphones could not be heard

[2] These stimuli were adopted from Goudbeek and Krahmer (2011), who created a slightly modified version of the stimuli used in Kellogg et al. (2007). Goudbeek and Krahmer used these stimuli for a secondary verbal memory task in a referential description experiment, and found that the dual task was cognitively demanding for most participants.

by their addressee, but that they had to pay attention to it nonetheless, since they would be asked about these sentences after the experiment as an attention check. Indeed, speakers were presented with 10 sentences at the end of the experiment, of which they had to select 5 that had occurred in the experiment. Speakers were also encouraged to pay attention to the dual task by way of a prize offer for the participant with the fewest errors. To keep the speaker aware of the addressee's needs, the addressee was allowed to ask the speaker clarification questions if anything remained unclear, but only after the speaker had finished the story.[3] After the experiment, speakers were asked informally about the difficulty of the secondary task, which was noted down by the experimenter.

The experiment was divided into two blocks, of which one contained the secondary task and the other did not, counterbalanced for order. Each block was preceded by two practice items. The experimenter was only present during the instructions and the practice trials. The experiment took about 25 minutes.

#### 5.5.1.4. Data coding

We transcribed all speakers' continuations of the third sentence, and coded all references to the target character (excluding possessives and reflexive pronouns). Since the target referent was referred to only once in the majority of the cases, we only analyzed the first subject reference. We excluded 34 cases in which the first subject did not refer to the target referent. In addition, we excluded 7 plural references, 3 indefinite references, 1 case in which the sentence presented over the headphones was repeated literally, and 1 missing case. In addition, there were 2 cases in which the referring expression was repaired. However, because the repair was of the same type in both cases (e.g., 'the man... uh the boy'), we kept these cases. In total, we excluded 46 trials (9.0%). The remaining 466 subject references were coded for the type of referring expression: either full noun phrase or pronoun.

#### 5.5.1.5. Design and statistical analyses

Crossing the two factors referent salience and cognitive load resulted in a 2 (speaker-salient, addressee-salient) x 2 (cognitive load, no cognitive load) within-participants design. Participants were assigned to one of four lists, each of which contained one version of a given item. The items were presented in a pseudo-random order, with at least one filler item between two consecutive experimental items.

[3] Speakers were still able to hear the addressee when wearing headphones.

We performed a logit mixed model analysis on the log odds for a pronoun (Jaeger, 2008). Referent salience and cognitive load were included as fixed factors, and participants and items as random factors. The fixed factors were centered to reduce collinearity. We attempted to fit a model with a full random effect structure. In case the model did not converge, we excluded random slopes with the lowest variance (as given by the non-converging model summary; Barr, Levy, Scheepers, & Tily, 2013). From the first converging model, we subsequently excluded random slopes that did not significantly contribute to model fit using log-likelihood ratio tests, with an $\alpha$-level of .20 (Baayen, Davidson, & Bates, 2008; Barr et al., 2013; Jaeger, 2011). Only the final model will be reported.

### *5.5.2. Results*

#### 5.5.2.1. Error rates

Six participants reported that they found the secondary task difficult, while 10 claimed to find it easy. Of 3 participants, no response was collected, and 13 had no strong opinion. The overall error rate in the secondary task (remembering the words BAL and DAL) was 9.0% (excluding 10 cases in which the response exceeded the time limit), suggesting that participants were at least paying attention to the secondary task. This rate was lower for those participants that reported to find the task easy (4.4%) than for those who claimed to find it difficult (17.6%). In the experimental items only, the error rate was 4.7%. Here, there was no significant effect of referent salience on the proportion of correct answers (logit mixed model with a by-participant random slope for referent salience: $\beta$ = -4.61; SE = 2.81; $p$ = .10), suggesting that the difficulty of the task did not differ between the two salience conditions.

Few errors were made in the attention check following the experiment (5.6%), suggesting that speakers were attending to the sentences presented over the headphones. In the addressee's task, i.e., selecting the correct picture pair from the booklet, the mean error rate was 3.5% (across both experimental and filler items), suggesting that this was relatively easy, and that the speakers' stories were generally interpretable. However, one addressee made 13 errors out of 36, which was mainly caused by the speaker incorrectly referring to the competing character. As noted above, such references were excluded from further analysis.

5.5.2.2. Proportion of pronouns

Figure 5.2 shows the proportion of pronoun references to the target character by referent salience and cognitive load condition. The final logit mixed model included random intercepts for participants ($s^2$ = 2.87) and items ($s^2$ = 0.15), as well as by-participant random slopes for referent salience ($s^2$ = 12.09) and cognitive load ($s^2$ = 3.04). We found a main effect of referent salience: Pronouns were more frequent when the referent was discourse salient only for the speaker (23.6%) than when it was discourse salient only for the addressee (8.3%), $\beta$ = 2.25; SE = 0.85; $p$ < .01. There was also a significant main effect of cognitive load, with slightly more pronouns in the cognitive load condition (17.2%) than in the no cognitive load condition (15.8%), $\beta$ = 1.37; SE = 0.56; $p$ < .05. However, these effects were qualified by a significant interaction between referent salience and cognitive load, $\beta$ = -2.76; SE = 0.95; $p$ < .01, suggesting that cognitive load affected pronoun use differently in the two salience conditions.

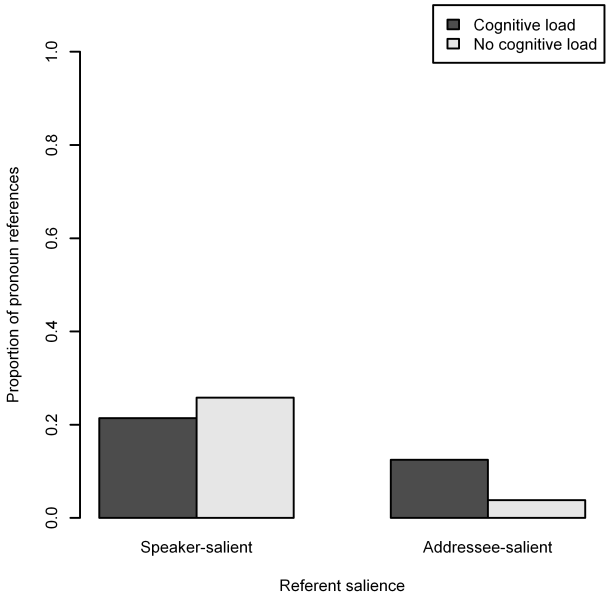


**Figure 5.2.** Proportion of pronoun references to the target character in the four conditions of Experiment 1.

To arrive at the pairwise comparisons for the interaction effect, we built separate models for the two levels of referent salience using the same procedure as described above. The final model for the speaker-salient condition included a by-participant

random slope for cognitive load, while the model for the addressee-salient condition included only by-participant and by-item random intercepts. The effect of cognitive load was not significant in the speaker-salient model, $\beta$ = 0.77; SE = 0.61; $p$ = .21, but marginally so in the addressee-salient model, with pronouns being more frequent in the cognitive load condition (12.5%) than in the no cognitive load condition (3.8%), $\beta$ = 1.35; SE = 0.77; $p$ = .08.

*5.5.3. Discussion*

The results of Experiment 1 show that pronouns were more frequent when the referent was discourse salient for the speaker but not for the addressee (i.e., it was the subject of the privileged context sentence), than when the referent was not salient for the speaker but was salient for the addressee. This suggests that speakers were more likely to use their own perspective to choose referring expressions than to take into account the addressee's perspective, even when they were not performing a secondary task. This is in line with Fukumura and Van Gompel (2012), who found that speakers tend to follow their own discourse model when there is privileged information.

More importantly, the results suggest that the presence of the secondary task increased the likelihood of pronoun use, at least in the condition in which the referent was not salient in the speaker's discourse model (addressee-salient condition). This does not support the hypothesis that speakers become more likely to fall back on their own discourse model when they experience an increased cognitive load. If that were the case, pronouns should have become less frequent under load when the referent was not salient for the speaker but salient for the addressee, and more frequent when the referent was salient for the speaker but not for the addressee. The results are also inconsistent with the claim that cognitive load on the part of the speaker decreases the accessibility of the mental representation of the referent in the speaker's discourse model (Arnold & Griffin, 2007). If that were the case, the execution of a dual task should have led to more specific expressions.

Our finding that the use of pronouns increases under load is compatible with the hypothesis that speakers under load are more likely to use expressions that are more economical for themselves (Hendriks et al., 2013). That is, an increased cognitive load may reduce the memory and attention resources required for determining that full noun phrases should be used to refer to entities that are not salient in the discourse, which may be a cognitively costly process. For referents that are salient in the

discourse, the use of pronouns is already in accordance with the referent's accessibility, which may explain why cognitive load does not increase pronoun use in this condition.

Since speakers appeared to be primarily making use of their own discourse model rather than calculating referent accessibility in the addressee's discourse model, the effect of cognitive load seems not to be due to difficulties in perspective taking. Still, given the relatively low overall proportion of pronouns in Experiment 1, it might be the case that speakers *were* taking into account the addressee's perspective, but not up till the level of calculating the referent's cognitive status for the addressee. This kind of detailed audience design might be cognitively too costly, even without an increased cognitive load (e.g., Bard et al., 2000; Brennan & Hanna, 2009; Horton & Gerrig, 2005). Therefore, speakers may just have increased the use of elaborate expressions to be as clear as possible for the addressee, as soon as they were aware of the fact that not all information was shared. This type of audience design might be more difficult under load, which could also explain the higher probability of pronoun use in the cognitive load condition.

To determine whether the increase in pronoun use under load in the addressee-salient condition in Experiment 1 is due to the speaker having difficulties in perspective taking, we conducted a second experiment in which all discourse information was shared between speaker and addressee. If the effect of cognitive load is due to the difference in perspective, changing privileged ground to common ground should cause this effect to disappear, since there is no need to increase the use of specific expressions when all information is shared (i.e., speaker's and addressee's discourse models match). If the effect of cognitive load is due to the speaker having difficulties in taking into account the salience of referents in the discourse, changing to common ground should not influence this effect, since referent salience remains the same for the speaker.

## 5.6. Experiment 2

### *5.6.1. Methods*

#### 5.6.1.1. Participants

Sixty-four students (44 female; mean age 22.3 years) from Tilburg University participated in the experiment for course credit. Half of them acted as speakers, the others acted as addressees. None of them participated in Experiment 1.

#### 5.6.1.2. Materials

We used the same experimental items as in Experiment 1. The only difference was that the speaker was not wearing headphones and that the second context sentence was presented over the computer speakers. As a result, both speaker and addressee had access to all discourse information.

#### 5.6.1.3. Procedure

The procedure was identical to that of Experiment 1, except that the speaker was not wearing headphones. As in Experiment 1, the speaker read aloud the first context sentence, which appeared below the first picture. After 5 s the second sentence was presented over the computer speakers. Speakers were told that they had to pay attention to this sentence, since they would be asked about these sentences after the experiment as an attention check. Next, the speaker completed the third context sentence based on the event shown in the picture. The addressee's task was the same as in Experiment 1, as was the dual task setup in the cognitive load condition. Again, the experiment was divided into two blocks, each preceded by two practice items. In contrast to Experiment 1, speakers were asked to indicate the difficulty of the secondary task on a 7-point Likert scale ('very easy' to 'very difficult').

#### 5.6.1.4. Data coding

The data coding procedure was the same as in Experiment 1. We excluded 1 case in which the first subject did not refer to the target referent and 1 plural reference (0.4%).[4] The remaining 510 subject references were coded for the type of referring expression: either full noun phrase or pronoun.

[4] We could speculate about reasons why we had much less missing data than in Experiment 1, but it is clear that the headphone manipulation in that experiment constituted a complicating factor that was not present in Experiment 2. See the discussion in Section 5.7.

5.6.1.5. Design and statistical analyses

Crossing the two factors referent salience and cognitive load resulted in a 2 (discourse salient, discourse non-salient) x 2 (cognitive load, no cognitive load) within-participants design. Participants were assigned to one of four lists, each of which contained one version of a given item. The items were presented in a pseudo-random order, with at least one filler item between two consecutive experimental items. Statistical analysis of the data was done in the same way as in Experiment 1.

### *5.6.2. Results*

5.6.2.1. Error rates

The mean reported difficulty of the dual task was 3.75 on a 7-point Likert scale, which seems to be in line with the informal judgments collected from Experiment 1. The overall error rate in the secondary task was 4.9%, and was lower for those participants that reported to find the task easy (a score of 3 or below; 2.0%) than for those who found it difficult (a score of 5 or above; 7.8%). In the experimental items only, the error rate was 4.7%, as in Experiment 1. Again, there was no significant effect of referent salience on the number of correct answers (logit mixed model with only by-participant and by-item random intercepts: $\beta = 0.80$; SE = 0.73; $p = .28$), suggesting that task difficulty did not differ between the salience conditions.

Slightly more errors than in Experiment 1 were made in the attention check following the experiment (8.8%), while the addressees made fewer errors on average in selecting the correct picture pair (1.1%). This might suggest that speakers did not need to concentrate as much on the sentences presented auditorily, since this was now shared information, and were able to produce clearer continuations.

5.6.2.2. Proportion of pronouns

Figure 5.3 shows the proportion of pronoun references to the target character by referent salience and cognitive load condition. The final logit mixed model included random intercepts for participants ($s^2 = 13.51$) and items ($s^2 = 0.26$), and a by-participant random slope for referent salience ($s^2 = 4.52$). We found a main effect of referent salience: Pronouns were more frequent when the referent was discourse salient (54.1%) than when it was not discourse salient (11.0%), $\beta = 6.24$; SE = 0.88; $p < .001$. There was also a significant main effect of cognitive load: More pronouns were used when speakers performed a dual task (34.5%) than when they did not (30.6%), $\beta$

= 0.76; SE = 0.39; $p < .05$. These effects were qualified by a marginally significant interaction, $\beta = -1.34$; SE = 0.77; $p = .08$.

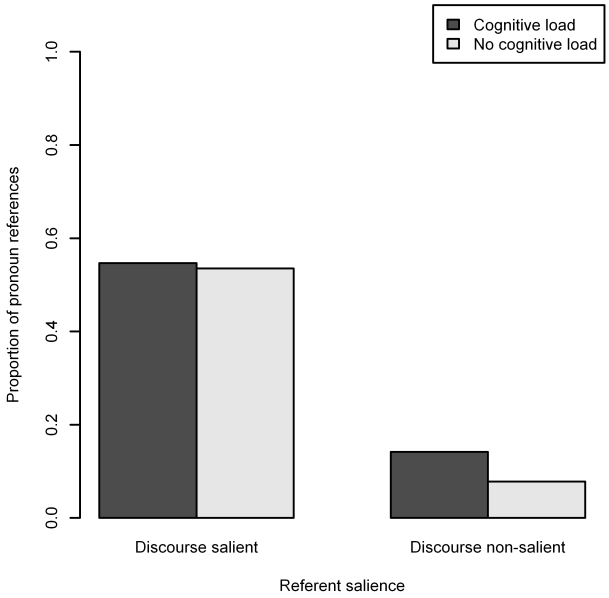


**Figure 5.3.** Proportion of pronoun references to the target character in the four conditions of Experiment 2

As in Experiment 1, we investigated this interaction further by building separate models for the two levels of referent salience. The final model for the discourse salient condition included only by-participant and by-item random intercepts, and the model for the discourse non-salient condition included only a by-participant random intercept. The effect of cognitive load was not significant in the discourse salient condition, $\beta = 0.11$; SE = 0.38; $p = .78$, but significant in the discourse non-salient condition, with pronouns being more frequent in the cognitive load condition (14.2%) than in the no cognitive load condition (7.8%), $\beta = 1.55$; SE = 0.67; $p < .05$.

*5.6.3. Discussion*

Experiment 2 was conducted to investigate whether the effect of cognitive load in Experiment 1 was due to speakers having difficulties in perspective taking, or whether it was due to difficulties in taking into account the referent's salience in the discourse. The results of Experiment 2 largely correspond to those of Experiment 1, except that speakers seemed to be more likely to produce pronouns, at least in the

condition where the referent was salient for the speaker.[5] This suggests that speakers were employing some kind of audience design in the presence of privileged information, resulting in more specific expressions. The observation that pronouns were more frequent in the cognitive load than in the no cognitive load condition is in line with the finding in Experiment 1 that adding cognitive load increases rather than decreases the probability of pronoun use. Again, this effect only seems to be present when the referent is not salient for the speaker. Thus, cognitive load seems to have the same effect when all discourse information is shared as when the salience of the referent differs for the speaker and the addressee.

Comparing the results of Experiment 1 and Experiment 2 indeed suggests that the effect of cognitive load does not differ across the two experiments. Combining the data of both experiments, we built a logit mixed model with referent salience (salient for the speaker, not salient for the speaker), cognitive load (cognitive load, no cognitive load), and experiment (shared context, privileged context) as fixed factors, and participants and items as random factors. Random slopes were only included when they improved model fit (again using log-likelihood ratio tests with an $\alpha$-level of .20).[6] The final model included random intercepts for participants ($s^2 = 8.73$) and items ($s^2 = 0.30$), as well as by-participant ($s^2 = 18.68$) and by-item ($s^2 = 1.32$) random slopes for referent salience.

There was a significant effect of referent salience, confirming that speakers were more likely to use pronouns when the referent was salient according to their own discourse model (39.0%) than when it was not salient (9.7%), $\beta = 5.53$; SE = 0.91; $p < .001$. The effect of cognitive load was also significant, confirming that speakers were more likely to use pronouns when they performed the secondary task (26.2%) than when they did not (23.6%), $\beta = 0.58$; SE = 0.29; $p < .05$. There was no main effect of experiment, $\beta = -1.09$; SE = 0.98; $p = .27$. However, there was a marginally significant interaction between referent salience and experiment, $\beta = -3.10$; SE = 1.71; $p = .07$, suggesting that the effect of salience was smaller in Experiment 1 (privileged context) than in Experiment 2 (shared context). There was also a significant interaction between referent salience and cognitive load, $\beta = -1.51$; SE = 0.59; $p < .05$, confirming

[5] Note that 'salient for the speaker' also means 'not salient for the addressee' in Experiment 1 (privileged context), but 'also salient for the addressee' in Experiment 2 (shared context). The reverse holds for 'not salient for the speaker'.

[6] The by-participants random slope for experiment was never included, since this factor was between-participants (i.e., participants in Experiment 1 were different from those in Experiment 2).

that the effect of cognitive load differed between the two salience conditions. Most importantly, the interaction between cognitive load and experiment, as well as the three-way interaction were not significant, $\beta$ = -0.24; SE = 0.61; $p$ = .69 and $\beta$ = -0.58; SE = 1.25; $p$ = .64, respectively. This indicates that the effect of cognitive load as well as the interaction between cognitive load and referent salience indeed did not differ between the two experiments.

Building separate models for the two levels of referent salience resulted in two models with only by-participant and by-item random intercepts. In the salient for the speaker condition, the effect of experiment was significant: Pronouns were less frequent when there was privileged context (23.6%) than when all discourse context was shared (54.1%), $\beta$ = -2.62; SE = 0.75; $p$ < .001. The effect of cognitive load was not significant, $\beta$ = -0.15; SE = 0.26; $p$ = .57, and neither was the interaction between cognitive load and experiment, $\beta$ = -0.53; SE = 0.53; $p$ = .32. In the not salient for the speaker condition, by contrast, the effect of experiment was not significant, $\beta$ = 0.41; SE = 2.64; $p$ = .88, while the effect of cognitive load was significant, with pronouns being more frequent in the cognitive load condition (13.4%) than in the no cognitive load condition (6.0%), $\beta$ = 1.42; SE = 0.55; $p$ < .01. The interaction between cognitive load and experiment was not significant, $\beta$ = 0.10; SE = 1.21; $p$ = .94.

These results indicate that the presence of privileged information causes the speaker to increase the use of full noun phrases, at least when the referent is salient for the speaker, in which case there is an initial preference for pronouns.[7] The effect of cognitive load seems to be independent of this type of audience design: Under load, speakers prefer to use less costly, more economic referring expressions for non-salient referents, both when there is privileged information and when all information is shared. This effect might therefore be primarily due to the speaker having difficulties in determining that full noun phrases should be used for referents that are not salient in the discourse.

[7] This is different from Fukumura and Van Gompel (2012), who found no interaction between referent salience and the presence of privileged information.

## 5.7. General discussion

### *5.7.1. Effects of cognitive load*

Two experiments investigated the influence of an increased cognitive load for the speaker on the choice of referring expressions. Speakers referred to both salient and non-salient entities (according to their own or their addressee's discourse model), either while performing a secondary task or not. Based on the literature, we formulated two alternative hypotheses concerning the impact of cognitive load on referential choice. Firstly, increased cognitive load may result in difficulties for the speaker in taking into account the addressee's needs. On the one hand, this may increase the speaker's tendency to choose referring expressions based on her own model of the discourse (e.g., Bard & Aylett, 2005; Horton & Keysar, 1996). On the other hand, increased cognitive load may cause speakers to resort to using more economical expressions (i.e., pronouns; Almor et al., 1999; Hendriks et al., 2008; 2013). Secondly, increased cognitive load may affect the speaker's own discourse model by decreasing the accessibility of referents therein, and hence lead to more elaborate expressions (Arnold et al., 2009; Arnold & Griffin, 2007), irrespective of salience. Alternatively, speakers may become less able to keep track of their own discourse model, causing their use of referring expressions to be less consistent (Arnold, 2010; Hendriks et al., 2013).

The results show that speakers under load become more likely to produce pronouns in a context that would normally require a full noun phrase. This effect of cognitive load does not seem to be related to the speaker's ability to take the perspective of the addressee, since speakers did not appear to calculate the referent's accessibility for the addressee even when they were not under load. In Experiment 1, as well as in Experiment 2, speakers generally used more pronouns when the referent was salient for them than when it was not salient for them, suggesting that they were basing their choice of referring expressions more on their own model of the discourse than on assumptions about their addressee's discourse model. Although they seemed to employ some kind of audience design by increasing the use of full noun phrases when there was privileged information (cf. Fukumura & Van Gompel, 2012), this was not harmed by the execution of a secondary task: We found a similar effect of cognitive load on the choice of referring expression in Experiment 2, in which speaker and addressee were assumed to have closely aligned discourse models, as in Experiment 1, in which referents that were not salient for the speaker were salient for the addressee.

In addition, our results do not provide support for the hypothesis that an increased cognitive load reduces the accessibility of referents in the speaker's own discourse model, since that would have resulted in an increase of full noun phrases. Hence, the present results suggest that although accessibility may be related to attention, it does not hold generally that less attentive speakers use more elaborate referring expressions. The increase in pronoun use for non-salient referents could be due to the speaker having trouble keeping track of the salience of the referent in her own discourse model, who therefore uses referring expressions less consistently. However, this does not explain why there is no decrease in the use of pronouns for salient referents. After all, if it becomes more difficult to keep track of referent salience, the choice of referring expressions should become less tied to the discourse salience of the referent (Arnold, 2010).

Our results support the hypothesis that cognitive load increases the use of more economical expressions. It is assumed that speakers prefer economical over elaborate expressions (Almor, 1999; Burzio, 1998; Levinson, 1987), and they may typically only use a full noun phrase when they refer to a character that is not salient in the discourse. Even though speakers do not seem to specifically keep track of the salience of the referent for the addressee, their own model of the discourse may serve as a proxy for that of their addressee (e.g., Bard & Aylett, 2005; Dell & Brown, 1991; Pickering & Garrod, 2004). Therefore, using this model to choose referring expressions can still be regarded as some kind of audience design. Thus, assuming that speakers themselves know what they are referring to, the production of full noun phrases may be inherently oriented towards an addressee. In fact, any expression that is more specific, and therefore more costly, than a pronoun could be considered an adaptation to a (hypothetical) addressee (Hendriks et al., 2013). Because this addressee-oriented process may be cognitively effortful, increasing the speaker's cognitive load may reduce this type of audience design. That is, when distracted by a secondary task, speakers may have fewer memory resources available that are needed to infer that a less salient referent should be referred to with a more elaborate expression. Hence, they are more likely to produce less costly expressions, such as pronouns.

One could also argue, however, that the effect of cognitive load is simply due to the fact that more elaborate expressions are more difficult to produce (i.e., an effect on lexical rather than conceptual representations). For example, with increased memory load, it may be more difficult to retrieve the richer semantic content of full noun phrases from memory, which results in the use of more semantically general

expressions such as pronouns (Almor, 1999). Although this explanation appears plausible, it does not explain why increased cognitive load only caused more pronouns when the referent was not salient for the speaker. If retrieving full noun phrases for production requires more cognitive resources, one would expect that using these would overall be more difficult when speakers experience increased cognitive load. Given that this did not appear to be the case in the present study, it seems more likely that the effect of cognitive load is related to assessing that referents that are not salient in the discourse should be referred to with more elaborate expressions, which may be an effortful process. The fact that cognitive load did not increase pronoun use when the referent was salient for the speaker may then be due to the coincidence with the preference to use pronouns anyway for such referents.

*5.7.2. Effects of dissociating the speaker's and addressee's perspectives*

As noted above, in addition to the use of full noun phrases for non-salient referents, another type of audience design, emerging from the comparison of the two experiments, is that speakers are more likely to use full noun phrases as soon as it is clear that the preceding discourse is not fully in common ground with their addressee. This may be evidence for a minimal, one-bit model of audience design (Galati & Brennan, 2010; see also Epley, Keysar, Van Boven, & Gilovich, 2004): Speakers use more specific referring expressions as soon as they are aware that not all information is shared, but irrespective of the actual accessibility of the referent for the addressee. This is in line with Fukumura and Van Gompel (2012), who found that while speakers were not taking into account their addressee's perspective in choosing referring expressions when the two perspectives were dissociated, they used slightly more pronouns in a condition in which all information was shared (37% vs. 33%), independently of whether the referent was salient or not. This suggests that speakers use more elaborate expressions when there is privileged information, even though they might run the risk of being overly specific.

Still, it is striking that the overall proportion of pronouns used in Experiment 1 was quite low. In their Experiment 2, for example, Fukumura and Van Gompel (2012) found higher rates of pronoun use in both the privileged, referent-salient condition (48%) and the privileged, referent-non-salient condition (18%) than we did in Experiment 1 (24% and 8%, respectively), while their results for the shared condition (referent-salient: 55%; referent-non-salient: 19%) were similar to our Experiment 2 (54% and 11%, respectively). Part of this difference could be explained by differences in the

linguistic materials. For example, while the referent mentioned in the second context sentence was referred to with a pronoun in Fukumura and Van Gompel's experiments, it was referred to with a full NP in our experiments, in accordance with the preferred way of referring to an entity previously mentioned as a direct object in centering theory (e.g., Brennan, 1995). The tendency to pronominalize the entity on a subsequent reference may be stronger when the referent had already been pronominalized. In addition, speakers may have been more likely to reuse the most recent referring expression, which could also have led to more pronouns in Fukumura and Van Gompel's experiments than in ours. However, this difference cannot explain why pronouns were also more frequent when the referent was not salient. Therefore, it seems that speakers in our study were employing minimal audience design more rigorously than in Fukumura and Van Gompel (2012).

One reason why speakers did not make the extra effort to calculate the accessibility of the referent in the addressee's discourse model may be that in the current experiment, as well as in Fukumura and Van Gompel's, references were never ambiguous, since the two characters always had a different gender. Therefore, not taking into account the addressee's perspective would probably not result in interpretation errors. However, in a situation in which it is clear that not taking into account the addressee's perspective *would* lead to interpretation errors, speakers may be more likely to adapt their choice of referring expressions to the knowledge of their addressee (e.g., Ferreira, Slevc, & Rogers, 2005; Horton & Keysar, 1996). In that case, increased cognitive load might make this perspective taking more difficult, and cause speakers to fall back on their own perspective.

The filler materials of Experiment 1 contained stories with characters of the same gender, and hence pronouns were ambiguous. Here, we indeed found more pronouns when the referent was salient for the addressee but not for the speaker (17; 33%) than when the referent was salient for the speaker but not for the addressee (7; 13%). This suggests that speakers were taking their addressee's perspective into account. However, cognitive load did not seem to cause speakers to use their own discourse model. Rather, a pattern similar to that in Figure 5.2 emerged, with more pronouns under load for referents that were not salient for the speaker. This may be another indication that the effect of cognitive load as manipulated here is independent of perspective taking.

The assumption in the above discussion has been that full noun phrases are more effortful for the speaker than pronouns (Almor, 1999), and are therefore dispreferred

unless they would aid the addressee's interpretation. However, it is conceivable, given the frequent overall use of full noun phrases in our experiments, that producing more elaborate referring expressions can also be helpful for the speaker herself. For example, names for salient entities may be more easily retrieved from the lexicon (Fukumura & Van Gompel, 2012). Alternatively, the use of full noun phrases may be related to the amount of conceptual information that needs to be retrieved to produce a referring expression (Engelhardt et al., 2006; Fukumura, Van Gompel, Harley, & Pickering, 2011; Fukumura & Van Gompel, 2012). Still, it is not clear how such speaker-internal explanations for the use of more elaborate expressions can account for the finding that the use of full noun phrases decreases under load.

Crucially, however, our results suggest that increased cognitive load does not harm the use of more elaborate expressions as a form of minimal audience design: The effect of cognitive load was the same in the presence of privileged information, in which case speakers often used full noun phrases, as when all information was shared. Hence, using more specific referring expressions when there is no full common ground may be relatively easy. Indeed, it has been argued that this type of audience design is cognitively not very demanding: Information that is available early or takes the form of a binary choice that can be assessed quickly (e.g., 'my addressee has heard this or not') is readily used in choosing referring expressions at little cost (Brennan & Hanna, 2009; Galati & Brennan, 2010). This might also explain the finding by Bard and Aylett (2005) that speakers only seemed to adapt their choice of referring expressions when they switched to a new addressee, since this is information that is easy to take into account (relative to, e.g., building a detailed model of someone else's knowledge). Thus, again, the decrease in the use of full noun phrases under load in our study does not seem to be related to difficulties in producing elaborate expressions per se, but rather to difficulties in determining *when* a full noun phrase should be used.

*5.7.3. Task-dependencies and individual differences*

It is important to note that the effect of cognitive load on referential choices may depend on the specific task and on what part of cognition is actually loaded. For example, in Arnold and Griffin (2007) and Fukumura et al. (2010), the number of referential competitors was varied, which led to divided attention to multiple possible referents. This might be a different kind of cognitive load than divided attention over multiple tasks. Thus, our results do not contest that referents competing for attention affect accessibility and therefore referential choices. However, our results are

inconsistent with the more general claim that accessibility and hence choice of referring expression is driven by speaker attention (Arnold & Griffin, 2007; Brennan, 1995). Rather, speakers may default to using pronouns when attention is led away from the discourse.

In a dual-task setup, the nature of the secondary task may also make a difference (e.g., whether it is visual or verbal; Baddeley & Hitch, 1974; Kellogg et al., 2007). For example, in the present experiments the use of a verbal secondary task may have especially hindered attention to the linguistic context. Other manipulations, such as a visual task (cf. Rosa & Arnold, 2011) or adding time pressure (cf. Horton & Keysar, 1996), possibly interfere more with activating non-linguistic representations or with perspective taking. In addition, it is conceivable that the artificial nature of the main, referential, task caused some additional load or encouraged the use of task-specific strategies. For example, the modality switches in the context sentences (reading aloud, then listening, then speaking) may have caused an extra increase in cognitive load. The same may hold for the unnatural dissociation of perspective using privileged information presented over headphones, which may therefore have caused a confound between perspective and cognitive load, although this cannot explain the effect of our independent manipulation of cognitive load using a secondary task. These issues need further research.

Finally, our results suggest that there was quite some individual variation as to how speakers' referring expressions were affected by the dual task. Although our cognitive load manipulation had an impact on referential choices, the secondary task appeared to be relatively easy for many participants. Informal inspection of the data suggested that participants who reported to have found the task difficult showed the clearest effects of cognitive load. One cause of individual differences in task difficulty could be the use of strategies for remembering the words BAL and DAL. Over the two experiments, two thirds of all participants reported to have used some kind of mnemonic (e.g., putting up one finger for BAL and two for DAL), although these were not always employed from the beginning. Since it is not yet clear how a heavier load might affect our results, this is a concern that should be taken up by future studies.

### 5.8. Conclusion

The study presented in this chapter has shown that speakers use more pronouns when they experience an increased cognitive load, at least when the referent is not salient for the speaker. We have suggested that this is due to difficulties in determining that a referent that is less salient in the discourse should be referred to with a more specific expression, which results in the production of more economical forms. We have not found support for the hypothesis that speakers under load are less able to take the addressee's discourse model into account, since speakers only appeared to use a cognitively undemanding form of perspective taking. Neither do our results support the hypothesis that cognitive load, at least in the form of the dual task used here, decreases the accessibility of referents in the speaker's discourse model or makes referential choices overall less consistent. Our results are in line with the view that speakers choose referring expressions based on assumptions about the referent's accessibility in the discourse, but make these assumptions primarily on the basis of their own model of the discourse.

# Chapter 6

# Discussion and conclusion

When a speaker produces an utterance, she has to select an entity to mention first as well as a linguistic expression to refer to that entity. These referential choices are believed to be influenced by the accessibility of the mental representation of the referent. The previous chapters have presented four studies, conducted in Dutch, that each focused on the interplay between two or more different factors, linguistic as well as non-linguistic and speaker-internal, that are assumed to affect this accessibility. In this final chapter, I will first summarize the main findings and formulate answers to the main research questions. Next, theoretical and methodological implications of these findings will be discussed and directions for future research will be set out. This chapter closes with a summary of the main conclusions.

### 6.1. Summary and answers to the research questions

The first research question, investigated in Chapters 2 and 3, was whether two non-linguistic saliency factors, visual foregrounding and perceptual animacy, affect the choice between pronouns and full noun phrases. The two story completion experiments presented in Chapter 2 showed that visually foregrounded characters were not more likely to be pronominalized than visually backgrounded characters. Thus, we found no evidence that visually foregrounding a referent leads to more attenuated referential forms. However, in Chapter 3 it was shown that perceptually animate entities were more likely to be pronominalized than perceptually inanimate ones. This suggests that at least some non-linguistic, perceptual saliency factors can influence the speaker's choice of referring expression.

The second research question, also investigated in Chapters 2 and 3, was how these non-linguistic factors interact with linguistic factors (discourse salience and lexical animacy). In both studies, the discourse salience of the referent (which was defined as whether the referent was mentioned in the preceding sentence, mostly as the subject) revealed itself as the strongest factor in determining the choice of referring expression. As shown in Chapter 3, perceptual animacy only affected pronoun use for referents that were not salient in the discourse. Similarly, in Chapter 2 the only situation in which a small tendency of an effect of visual salience on the choice of referring expression was found was when the referent was non-salient in the discourse. For discourse-salient referents, pronouns were highly preferred in both studies, and this left little room for either animacy or visual salience to take effect. Another linguistic

factor that we investigated was lexical animacy. This factor was shown to be overruled by both discourse salience and perceptual animacy in the choice of referring expression.

The third research question investigated in these chapters was whether effects of linguistic and non-linguistic factors are different for the choice of referring expression on the one hand, and the choice of referent on the other. Even though both linguistic and non-linguistic saliency factors are assumed to underlie the accessibility of mental representations, and hence referential choices, we indeed found that the factors investigated here affected these referential choices differently. Notably, while we found no convincing evidence for visually salient entities being pronominalized more often, visual salience did affect which referent speakers were more likely to mention first, as the subject of their utterance. Similarly, entities that had animate names (but were not necessarily perceived as animate) had no increased probability of pronoun use when the entity's discourse salience was controlled for, but they were more likely candidates for the subject role than entities with inanimate names. The opposite pattern was observed for perceptual animacy, where entities that were perceived as animate were more likely to be pronominalized than entities perceived as inanimate, although they were not more likely to be mentioned as the subject of the sentence. Only the referent's salience in the discourse, as measured by whether participants had mentioned the referent in the previous utterance, was found to affect both referential choices. As for this factor, however, the experiments described in Chapter 2 suggested that there may be a difference between referents that are explicitly introduced as the topic of the discourse and referents that are local topics. The former influenced first mention, whereas the latter did not.

Furthermore, it might not be the same factors that affect the choice between a pronoun and a full noun phrase and that affect the choice between other referential forms, such as that between full and reduced pronouns. Given that animacy appears to affect pronominalization, our next research question was whether animacy also affects the choice between full and reduced pronouns. Chapter 4 confirmed that in Dutch, animate referents are indeed more likely to be pronominalized than inanimate referents, but also showed that they are more likely to be referred to with full pronouns than with reduced pronouns. This is not in accordance with the claim that the accessibility scale maps on a hierarchy of referential forms that goes from less to more attenuated expressions. It was suggested that the choice between full and

reduced pronouns might be driven more by the importance of the information conveyed by the referring expression.

Animacy effects in Dutch may also arise because pronouns encode gender information, while many native speakers are uncertain about what grammatical gender inanimate nouns should be assigned to. Hence, this might be another reason why pronouns are avoided for inanimate referents. Therefore, if in addition to accessibility, gender avoidance plays a role in the effect of animacy on pronoun use, the effect should be different in language varieties with and without a tendency to avoid gendered pronouns for inanimate antecedents. The study described in Chapter 4 compared pronoun production of speakers of Netherlandic Dutch, who generally lack intuitions about masculine and feminine grammatical gender of inanimate nouns, with pronoun production of speakers of Belgian Dutch, for whom these intuitions are typically strong. A sentence completion experiment was conducted with nominal antecedents that were pretested for these intuitions. The results of the experiment showed that the effect of animacy on the choice between pronouns and full noun phrases in Dutch cannot be explained by a gender avoidance strategy only: Speakers of both varieties were less likely to pronominalize inanimate referents, although the overall rate of pronoun use was higher for the Belgian Dutch speakers than for the Netherlandic Dutch speakers. We concluded that the animacy effect in Dutch is likely to be driven by the higher conceptual accessibility of animate referents.

The final research question, addressed by Chapter 5, was whether it is the accessibility of representations in the mind of the addressee – as estimated by the speaker – that matters for the choice of referring expression, or whether the accessibility of representations in the speaker's own mind plays a role as well. Hypothesizing that a referent's accessibility for the speaker decreases when she is experiencing an increased cognitive load, the two story completion experiments in this chapter tested the effect of the addition of a secondary task for the speaker on the use of pronouns, in interaction with the referent's salience in the discourse (either from the speaker's or from the addressee's perspective). An increased cognitive load for the speaker did not decrease the use of pronouns. Rather, speakers under load were *more* likely to use pronouns, at least for referents that were not salient in the discourse. This suggests that accessibility is not necessarily determined by how much attention the speaker herself has allocated to the referent. However, speakers did not seem to keep track of the referent's accessibility from the addressee's perspective

either. These findings were taken as evidence that speakers are producing referring expressions based on their assumptions about the referent's accessibility in the discourse, but make these assumptions based on their own model of the discourse rather than on that of their addressee. When under increased cognitive load, speakers have trouble taking into account the referent's salience in this model, and fall back on the egocentric preference of producing more economical expressions.

## 6.2. Theoretical implications

This section discusses implications of the research presented in this dissertation for existing theories of reference. First, Section 6.2.1 restates the different approaches to accessibility in research on the choice of referent for first mention and in research on the choice of referring expressions. Next, Section 6.2.2 indicates where the present research supports existing theories of reference and where it diverges from them. A tentative proposal of how seemingly contradictive effects of accessibility-related factors can be unified in a single model is presented in Section 6.2.3. Finally, Section 6.2.4 briefly discusses implications for computational models of referring expression generation.

### *6.2.1. The opposition between the choice of referent and the choice of referring expression*

This dissertation concerned the accessibility of mental representations of referents, and its effect on referential choices in spoken language production. Accessibility has been taken to be the single notion that makes it more likely that referents are mentioned first as well as that they are pronominalized. The problem is that the notion of accessibility has been approached somewhat differently in research on these two referential choices. On the one hand, the choice of what to mention first has generally been considered a speaker-internal process of activating and selecting (lexical) concepts (Levelt, 1989). A concept is assumed to have a high activation, for example, when it has been mentioned recently (Ferreira & Yoshita, 2003), when it is triggered by what the speaker is perceptually attending to (Gleitman, January, Nappa, & Trueswell, 2007), or when it is inherently salient (e.g., animate; Branigan, Pickering, & Tanaka, 2008). As soon as this activation reaches a certain threshold, the concept is passed down to the level of linguistic formulation, where it is assigned a prominent syntactic position (either directly, or through the assignment of grammatical functions or information structure).

The selection of a certain type of expression to refer to that concept, however, has traditionally been seen as a highly addressee-oriented process (e.g., Ariel, 1990; Clark, Schreuder, & Buttrick, 1983; Gundel, Hedberg, & Zacharski, 1993). A speaker chooses a referring expression such that her addressee is able to readily resolve what is being referred to. Therefore, the speaker has to make assumptions about what information she has in common ground with her addressee. An important source of common ground is the preceding discourse, which can generally be assumed to be accessible for both speaker and addressee (Clark, 1996). In sum, whereas the selection of concepts for next mention is assumed to be driven largely by the accessibility of representations for the speaker, the selection of a referring expression is assumed to be dependent in the first place on the speaker's assumptions about the accessibility of representations for the addressee.

*6.2.2. Implications for theories of reference*

The aim of this dissertation was to investigate whether the speaker-internal and non-linguistic factors that have been found to affect the choice of referent for first mention also play a role in the choice of referring expression. The findings of the studies presented here paint a mixed picture with respect to this central question.

First and foremost, out of the factors tested here, both a non-linguistic factor, namely animacy, and a speaker-internal factor, namely cognitive load, were indeed shown to have an impact on the choice of referring expression. As for animacy, the observation that animate entities were more likely to be pronominalized than inanimate entities (Chapters 3 and 4) corroborates findings by Fukumura and Van Gompel (2011). In Chapter 3, it was additionally shown that this animacy effect is likely to be driven by the conceptualization of the referent in the (perceptual) context, rather than by the lexical semantics of the antecedent. This implies that the choice of referring expression is not only determined by structural linguistic factors, such as grammatical function and topichood, as some researchers have claimed (e.g., Fukumura & Van Gompel, 2010; Kehler, Kertz, Rohde, & Elman, 2008; Stevenson, 2002), but can also be influenced by non-linguistic factors. Crucially, the finding that the choice of referent for first mention was only influenced by the referent's lexical animacy suggests that animacy has different effects on different referential choices.

With respect to the question whether referential choices are mainly driven by addressee-oriented factors or by speaker-internal constraints, the present research also suggests a difference between the choice of referent and the choice of referring

expression. This is in line with how these referential choices are described in most accounts of reference. For example, studies on the linearization of constituents have suggested that the choice of referent for first mention is determined by the degree to which the speaker attends to the referent (e.g., Gleitman et al., 2007; Tomlin, 1997). If speaker attention determines accessibility, one would expect that it also affects the choice of referring expression, as suggested by Arnold and Griffin (2007). However, the findings in Chapter 5 suggest that entities receiving less attention from the speaker are not necessarily referred to with less attenuated expressions.[1] This supports the view that the choice of referring expression instead depends on a model of the discourse that the speaker maintains to be able to select referring expressions that a (hypothetical) addressee can interpret correctly (e.g., Hendriks, Koster, & Hoeks, 2013). This model at least includes information about accessibility based on (local) discourse salience. Using this model to choose referring expressions is a type of audience design in the sense that the speaker makes assumptions about how accessible referents are in the discourse. However, the present research has provided additional evidence that this is not the same as explicitly calculating the accessibility of referents from the perspective of the addressee, because the speaker bases the model on her own knowledge of the discourse (cf. Fukumura & Van Gompel, 2012; Hendriks et al., 2013).

Thus, the findings for both animacy and cognitive load suggest that even though non-linguistic and speaker-internal factors that are associated with the choice of referent for first mention may also influence the choice of referring expression, they do not have the same effects in both referential choices, in line with Kehler et al.'s (2008) Bayesian approach. This is further supported by the results of Chapter 2, where the referent's visual salience affected the former, but not the latter.

Second, in all studies except Chapter 4, the discourse salience (grammatical function or topicality) of the antecedent still appeared to be the most important factor in the choice of referring expression (but in Chapter 4, grammatical function did affect the choice between a pronoun and a demonstrative). For example, both perceptual animacy and cognitive load only affected pronoun use for referents that were not salient in the discourse, suggesting that these factors were overruled by discourse salience. These findings are in line with traditional theories of reference, such as Ariel

[1] This is based on the assumption that speakers attend more to referents when they are focused on the referential task than when they are cognitively loaded with a secondary memory task. Whether speakers really refer differently to entities that are in their focus of attention is a question that still needs to be empirically tested, perhaps using eye-tracking studies.

(1990) and Gundel et al. (1993), as well as with centering theory (Grosz, Joshi, & Weinstein, 1995), which assign the greatest importance in determining the accessibility of a referent to its local discourse salience. They are also in line with psycholinguistic studies showing that the influence of the non-linguistic context becomes smaller when salience in the linguistic context is more compelling (e.g., Fukumura, Van Gompel, & Pickering, 2010).

Finally, especially Chapter 4 contributed to the evidence that the choice of referring expression is not a unified phenomenon, as the widely used accessibility scales such as those proposed by Ariel (1990), Givón (1983) and Gundel et al. (1993) make it appear. Which type of referring expression is chosen is not just a function of its association with a particular accessibility status; multiple factors may play a role, which each may have different effects on different expressions. This is in line with a form-specific multiple constraints approach to reference (Kaiser & Trueswell, 2008), according to which there are differences between specific referential forms in the degree to which they are sensitive to certain factors.

*6.2.3. A tentative proposal for a unified account*

The overall picture emerging from the results presented in this dissertation confirms the prediction that accessibility-related factors affect the choice of referent and the choice of referential form differently (cf. Kehler et al., 2008). The proposal made here is that, on the one hand, the selection of a referent for first mention may be driven by the speaker's *global conceptualization of the discourse*, perhaps in the form of a mental model (Johnson-Laird, 1983; Van Nice & Dietrich, 2003a). In a narrative discourse, this model may contain information about, for example, who or what the main or most important character is and about who does what to whom and why (Griffin & Bock, 2000). It may also contain contextual information about the event being described and the communicative situation (e.g., 'who is my addressee?'; Galati & Brennan, 2010). Who or what is conceptualized as the main character may in turn be influenced by whether the character was linguistically introduced as such (e.g., *Once upon a time there was a little dragon...*), by whether it is perceptually salient (e.g., visually foregrounded), or by whether it is conceptually salient (e.g., human agents are more likely to be main characters than stones or snowflakes). These factors make a referent accessible in a global representation of the discourse, although they can be in conflict, for example when the discourse topic is not a human agent (see also Montag & MacDonald, 2013 for interactions between animacy, visual salience and linguistic context).

The choice of a particular referring expression, on the other hand, often depends on a more *local model of the discourse*, involving primarily linguistic factors (cf. Grosz, 1977; Grosz et al., 1995). For example, pronouns are likely to be used when the antecedent was the subject or the topic of the directly preceding sentence. However, this does not mean that there is no room for other factors to play a role. Notably, we have seen in Chapter 3 that the perceived animacy of the referent influences the choice of referring expression. Crucially, this effect was independent of how the referent was lexically introduced. Thus, it seems that speakers are more likely to pronominalize those entities that they have conceptualized as animate or human based on non-linguistic information. This might be part of their mental model of the event, in which entities higher in animacy are more accessible, for example because they are more likely to be main characters and more important for the event to be described.

The distinction between a local and a global representation of the discourse in referential choices could be visualized as in Figure 6.1. As shown by the arrows, both models influence both the choice of referent as well as the choice of referring expression. However, the amount of influence differs, in accordance with the finding that accessibility-related factors affect both referential choices differently. This could be represented by giving weight values to the connections. For example, given that referent accessibility in the global discourse model is probably more influential in the choice of referent, the weight of the corresponding connection is likely to be higher than that of the arrow departing from the local discourse model. Conversely, given that local discourse factors strongly affect the choice of certain types of referring expression, the link from the local discourse model to particular referring expression types is likely to have a higher weight than the link from the global discourse model.

In addition, different types of referring expression, such as full and reduced pronouns, may also differ in their sensitivity to both local and global discourse salience, in line with Kaiser and Trueswell's (2008) form-specific multiple-constraints approach to accessibility. Kaiser and Trueswell (2008) argue that entities can be accessible on different levels, distinguishing between a local, syntacto-semantic and a global, mental model level of accessibility. Choices for different types of referring expression are sensitive to different levels of accessibility. Also the same type of referring expression may be sensitive to both global and local sources of accessibility to differing degrees. Again, global factors affecting the choice of referring expression, such as the importance of entities in the event to be described, may run counter to a classical accessibility account (e.g., Ariel, 1990), such that referents that should count

as more accessible are actually preferred to be referred to with the less reduced forms. The present proposal extends this account to the choice of referent for first mention: What is most likely to be mentioned next may be affected by both levels of accessibility, but the degree to which these levels are involved differs, and is not the same as in the choice of a particular referring expression.

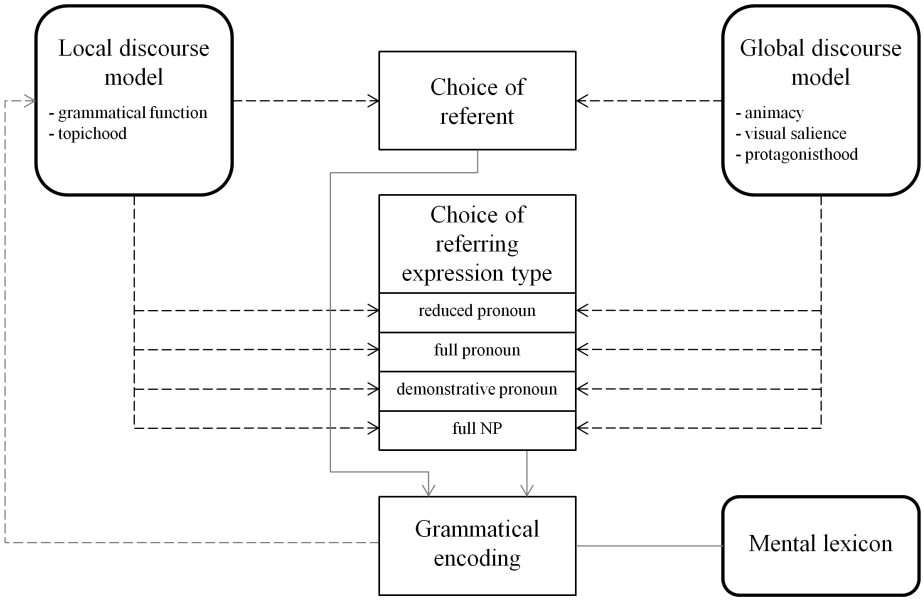


**Figure 6.1.** Tentative proposal of a unified model of effects of accessibility from local and global discourse models on the choice of referent and the choice of referring expression type. Included are those factors that the present research has identified as possibly affecting accessibility in the respective discourse models. Solid lines mean "provides input for"; dashed lines mean "influences".

How this proposal might be integrated into a general model of language production (e.g., Levelt, 1989) is shown by the grey colored connections in Figure 6.1. As soon as a speaker has chosen a referent for next mention, the referent is probably assigned a grammatical function in the grammatical encoding stage (Levelt, 1989). In this stage, the referent will also be assigned the selected type of referring expression. The lexical item required to express the selected referent is retrieved from the mental lexicon. The ease of retrieval of lexical items (lexical accessibility) may also influence grammatical encoding, e.g., whether the item is produced as the subject of the sentence. The assignment of grammatical functions in the grammatical encoding stage may in turn

result in an update of the local discourse model, with more prominent grammatical functions increasing the referent's accessibility in this model.

The influence of factors affecting the global mental model will often be in line with local factors. For example, the global discourse topic is likely to also be the topic of individual utterances (e.g., Givón, 1983). Global factors may also affect the local discourse model indirectly: A globally accessible entity is more likely to be mentioned first in a particular utterance, which in turn increases its local accessibility (via the connection between grammatical encoding and the local discourse model in Figure 6.1). This will affect the referential form with which this entity is referred to in the next utterance. In other cases, the two types of factors may be in conflict, such as when the local topic is inanimate (see Chapter 3), or when the local topic is not the same as the global discourse topic or the protagonist (e.g., Poesio, Stevenson, Di Eugenio, & Hitzeman, 2004; Van Vliet, 2008). In sum, although local discourse factors might dominate in the choice to pronominalize (e.g., the effect of perceptual animacy in Chapter 3 was not present when the referent was highly discourse salient), more global conceptual factors can still play a role.

The distinction between local and global accessibility may partly overlap with the distinction between derived and inherent accessibility (Prat-Sala & Branigan, 2000): Inherent properties of referents such as animacy are likely to influence the global model of the discourse, while properties derived from context such as topicality may affect local discourse more. Crucially, however, Chapter 3 has suggested that this distinction is not always easy to make, since inherent properties such as animacy may also be influenced by the context. After all, it holds for animacy as well as for other factors that what determines accessibility is not the properties of entities per se, but how the entities are *conceptualized* (Jackendoff, 2002). Furthermore, our results do not seem to support an account in which the accessibility of referents is equaled to their predictability in context (e.g., Arnold, 2001; 2008; Givón, 1983). This is because if effects of accessibility could be explained by predictability alone, we should have found that what is most predictable, i.e., what is most likely to be mentioned next, is also most likely to be referred to with more attenuated expressions. The results of the present research indicate that this is not necessarily the case, in line with findings from other studies (e.g., Fukumura & Van Gompel, 2010; Kaiser, Li, & Holsinger, 2011; Rohde, 2008).

Of course, the model presented here is highly sketchy, and in its current state not very useful to make specific predictions about how a certain referential choice will be

influenced by different factors. Clearly, further research has to be conducted to investigate the exact role of various global and local factors in referential choices.

*6.2.4. Implications for computational models of referring expression generation*

The present research also has implications for computational models of referring expression generation. Most existing models either produce references not embedded in a linguistic context or implement some basic account of linguistic salience (e.g., Dale & Reiter, 1995; Krahmer & Theune, 2002). For human-like generation of referring expressions, future models should also take into account influences of factors that go beyond the local discourse context (e.g., McCoy & Strube, 1999). This could perhaps be done by incorporating a model of global accessibility, which takes input from properties of the physical context, intrinsic properties of referents and narrative structure, for example. The local and global sources of accessibility may influence the likelihood that a particular expression is chosen probabilistically, with probabilities adjusted for each type of expression.

## 6.3. Methodological implications

Most psycholinguistic research on the production of referring expressions in discourse has made use of story completion (or continuation) experiments. Chapters 2, 4 and 5 of this dissertation have followed this tradition (although the study presented in Chapter 4 was more constrained, as only the referring expression needed to be filled in). While many studies have used written completion tasks (e.g., Anderson, Garrod, & Sanford, 1983; Fukumura & Van Gompel, 2011; Kaiser & Trueswell, 2004; Stevenson, Crawley, & Kleinman, 1994; Vonk, Hustinx, & Simons, 1992), we focused on spoken language production only (cf. Arnold, 2001; Arnold & Griffin, 2007; Fukumura & Van Gompel, 2012; Fukumura et al., 2010; Kaiser et al., 2011). Written tasks might not elicit the same kind of results as would spoken language production experiments. For example, when people speak, they have less opportunity to reflect on what they are saying than when they write. Hence, the spoken modality might give more insight into people's initial, automatic linguistic choices (Arnold, 2001).

One concern with story completion studies is their ecological validity, i.e., they often lack sufficient resemblance to naturalistic communicative situations. We have tried to tackle this issue in a number of ways. Firstly, many story completion tasks are not

embedded within a visual context, while many everyday communicative situations are. If visual context is used, it is often highly artificial or cartoonish. The studies presented in Chapters 2 and 5 have tried to remedy this by using stimuli consisting of photographs of real people. However, a problem with naturalistic scenes is that they quickly become cluttered, which may influence reference production (e.g., Coco & Keller, 2009; Koolen, Krahmer, & Swerts, 2013). This increases the number of factors that may affect referential choices, which makes analysis more difficult. The solution in the stimulus material used in the chapters mentioned above was to include only a minimal number of people and objects photographed against a neutral background.

Secondly, the visual context in the experiments described in Chapters 2 and 5 remained available during story continuation, and the same was true for the linguistic context in Chapter 4. While this is not necessarily unrealistic, the physical context could serve as something to hold on to for the speaker, and this might affect language production (Van Nice & Dietrich, 2003b). A setup as in Chapter 3, in which people speak from memory when retelling an event, might mimic everyday situations in which people do not talk about things that are immediately present (cf. Christianson & Ferreira, 2005). It might be that with such a setup, one is most likely to find interactions between different kinds of factors, since they are all processed at more or less the same time (i.e., everything has to be kept in memory; Van Nice & Dietrich, 2003b). However, a drawback of a retelling task is that as a researcher, one has little control over the context of an utterance. For example, in the experiments in Chapter 3 it was not possible to manipulate the local discourse salience of the target referent, as this depended on the participant's own previous utterances. Using a constrained (visual or linguistic) context allows for more control over what people refer to.

A further aspect of the naturalness of a story completion experiment is the goal of the task and the role of the addressee. In the studies presented in Chapters 2 and 4, as is common in other studies, participants were producing their utterances in a lab, in front of a computer screen, with no one else around. Utterances produced in such an experimental setting are likely to differ from those produced in more everyday communicative situations, in which speakers have the intention to convey a message to someone with a particular goal in mind (see also Montag & MacDonald, 2013). Indeed, there is evidence that the actual presence of an addressee, as well as his degree of engagement, makes a difference in the use of referring expressions (e.g., Arts, 2004; Kantola & Van Gompel, 2011; Rosa, Finch, Bergeson, & Arnold, 2013). At the same time, there is also evidence that even when no addressee is present, speakers

show linguistic behavior that would otherwise be considered a form of audience design (e.g., Koolen, Gatt, Goudbeek, & Krahmer, 2011; Van der Wege, 2009). Therefore, it remains an open question how much of the effects attributed to the presence of an addressee are actually speaker-internal.

A final advantage of experimental tasks that allow the participants to talk freely (with or without an addressee) over completion tasks is that they do not suffer from shifts in modality or breaks in the speech flow. Obviously, it is more naturalistic to have participants tell their own stories than to first have them listen to or read (aloud) part of the story and then have them switch to speaking from their own imagination, which also involves a switch from comprehension to production. Again, however, this latter setup may be necessary to systematically investigate effects of the linguistic context. Therefore, we would not argue for or against one or the other method, as they all have their merits and drawbacks. However, we would argue that the nature of the task is something that should be taken into account when discussing the outcomes of any experiment. Concerning the studies conducted for the present research, it would be interesting to conduct similar experiments in highly naturalistic, unconstrained contexts for comparison. These may also include other types of discourse than narratives. In addition, experimental research might be complemented with data from corpus studies or other naturally collected data to see whether the results would generalize to language use outside the lab.

### 6.4. Suggestions for future research

In this section, some remaining questions will be discussed briefly. One question is what the locus is of accessibility effects in the human mind. Accessibility is believed to be a property of representations in memory (e.g., Ariel, 1990; Bock & Warren, 1985), but the exact relation between reference and memory often remains implicit (cf. Chafe, 1994; Van Nice & Dietrich, 2003b). One idea is that the degree of accessibility relates to whether a referent is represented in long-term or in short-term memory (e.g., Gundel et al., 1993): Referents are more accessible when they are in short-term than in long-term memory. Alternatively, accessibility may only be a property of working memory. Van Rij (2012) proposes a computational model of the production and comprehension of pronouns, in which accessibility is modeled as a combination of base and spreading activation of entities in working memory. The entity with the highest activation is

pronominalized. A low working memory capacity causes fewer spreading activation, affecting the speaker's ability to determine the referent's salience. However, if referents can be accessible on multiple levels (e.g., global vs. local), as suggested above, it is not clear how these different levels of accessibility would connect to a single memory representation. It is therefore an open question whether it is possible for these different levels to be subsumed under a single notion of accessibility, or whether they actually represent different cognitive processes. Since the role of accessibility in reference production has been studied mostly by looking at the output (i.e., the referring expressions produced), no direct measures of accessibility are available. Future research in this field may therefore benefit from studies that tap into the cognitive processes underlying reference production themselves, for example by using eye-tracking or manipulating memory load.

Another question, already discussed to some degree, is to what degree referential choices are shaped by speaker-internal processes and to what degree they are addressee-oriented (Arnold, 2008). The results of Chapter 5 suggested that accessibility resulting from local discourse salience is not necessarily related to the amount of attention the speaker herself has allocated to the referent. However, it might still be the case that accessibility resulting from non-linguistic factors such as visual salience and animacy *is* speaker-internal, for instance because these factors are not dependent on a local model of the preceding discourse. To investigate this, one could test the effect of an increased cognitive load on the influence of these factors. For example, a speaker under load might produce more specific referring expressions for animate entities than without cognitive load, due to a decrease in the accessibility of the referent in her own memory.

Third, this dissertation has been concerned mainly with the accessibility of concepts, not of lexical items. It has been suggested, however, that the choice of referent for first mention is at least partly driven by the ease of retrieval of lemma representations (e.g., Gleitman et al., 2007). This seems to be supported by our finding in Chapter 3 that lexical but not contextually driven animacy affected mention of the referent in subject position. However, given that visually salient entities are also more likely to be mentioned first (see Chapter 2), such an account would suppose a direct link between perceptual properties of referents and linguistic representations, without the intervention of concepts. Whether such a link is cognitively plausible should be researched further. Furthermore, pronouns may be more lexically accessible than full noun phrases, being semantically and phonologically reduced. Hence, pronouns may

be selected most easily in the choice of referring expression, in accordance with the idea that the use of pronouns to refer to an entity is default, the use of more specific expressions being triggered by addressee-oriented requirements (Hendriks et al., 2013). As a result of that, a lexical accessibility account would also predict that pronouns are more likely to be mentioned first in an utterance. Such effects have not yet been taken into account in the present studies.

Finally, the present dissertation has mainly focused on the use of third person singular personal pronouns versus more specific types of referring expression (e.g., full NPs) in references to persons and objects. The range of possible referents and referring expressions is, however, much larger than that. To name a few, one could investigate references to spatial locations or points in time (e.g., *here*, *there*, *tomorrow*), to events and other abstract entities (e.g., *yesterday's thunderstorm*, *my dream*), to substances (e.g., *the mud*) to parts of objects (e.g., *the sheep's nose*) or to sounds (e.g., *that noise*); one could also investigate the use of different types of definite and indefinite descriptions (e.g., with specific or generic referents), deictics (e.g., *this*, *that*), second person pronouns (e.g., *you*), plural expressions or zero anaphora.

In addition, as was briefly hinted at in Chapter 4, effects of accessibility can be different in other dimensions than the choice of referential form, such as acoustic reduction, disfluency and gesture. Recent work in psycholinguistics has already started to explore some of these directions (e.g., Arnold & Tanenhaus, 2011; Gullberg, 2006; Hoetjes, Koolen, Goudbeek, Krahmer, & Swerts, 2011; Kaiser et al., 2011; Watson, Arnold, & Tanenhaus, 2008). Furthermore, while most of the present research has been conducted using university students (with the exception of Chapter 4), reference production should also be studied in other groups of participants. Work that has already been done includes research on young children (e.g., Hendriks et al., 2013; Matthews, Lieven, Theakston, & Tomasello, 2006), elderly people (e.g., Hendriks et al., 2013), Alzheimer patients (Almor, Kempler, MacDonald, Andersen, & Tyler, 1999) and people with autism (e.g., Arnold, Benetto, & Diehl, 2009). Future work might hopefully benefit from the issues brought forward in the present dissertation. Finally, computational models of reference production might be developed that incorporate insights emerging from a broader view on referential choices.

## 6.5. Conclusion

This dissertation has been concerned with the question how speakers choose what they will mention first in an utterance and how they will refer to it, as well as with the role the accessibility of mental representations plays therein. Traditionally, effects of accessibility on the choice of a referent for first mention have been ascribed to the amount of attention the speaker has allocated to a referent, which may be influenced by linguistic, perceptual or conceptual factors. By contrast, effects of accessibility on the choice of referential form have been explained as the speaker's effort to choose expressions in accordance with the information she has in common ground with her addressee, which is largely determined by linguistic properties of the directly preceding discourse. The studies presented here have provided evidence for the influence of both a non-linguistic factor (perceptual animacy) and a speaker-internal factor (cognitive load) on the choice of referring expression, interacting with linguistic factors. At the same time, they have also made it clear that there are indeed differences in the effects of accessibility on different referential choices. This has consequences for our understanding of the notion of accessibility, since it suggests that there are multiple types or levels of accessibility, to which different language production processes can be sensitive in different ways. These findings must be incorporated into both theoretical and computational models of reference. More generally, they may provide new insights in how we think about human cognition and communication.

# References


Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review, 106*(4), 748-765.

Almor, A., Kempler, D., MacDonald, M. C., Andersen, E. S., & Tyler, L. K. (1999). Why Do Alzheimer Patients Have Difficulty with Pronouns? Working Memory, Semantics, and Reference in Comprehension and Production in Alzheimer's Disease. *Brain and Language, 67*, 202-227.

Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *The Quarterly Journal of Experimental Psychology, 35*(3), 427-440.

Anderson, G. D. S. (1997). On "Animacy Maximization" in Fox (Mesquakie). *International Journal of American Linguistics, 63*(2), 227-247.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.

Ariel, M. (1998). The linguistic status of the "here and now". *Cognitive Linguistics, 9*(3), 189-237.

Ariel, M. (2001). Accessibility theory: an overview. In T. Sanders, J. Schilperoord & W. Spooren (Eds.), *Text Representation: Linguistic and psycholinguistic aspects* (pp. 29-87). Amsterdam/Philadelphia: John Benjamins.

Arnold, J. E. (1998). *Reference form and discourse patterns.* PhD dissertation, Stanford University, Palo Alto, CA. Retrieved from http://www.unc.edu/~jarnold/papers/diss/fulldissertation.doc

Arnold, J. E. (2001). The Effect of Thematic Roles on Pronoun Use and Frequency of Reference Continuation. *Discourse Processes, 31*(2), 137-162.

Arnold, J. E. (2008). Reference Production: Production-internal and Addressee-oriented processes. *Language and Cognitive Processes*, *23*(4), 495-527.

Arnold, J. E. (2010). How Speakers Refer: The Role of Accessibility. *Language and Linguistics Compass*, *4*(4), 187-203.

Arnold, J. E., & Griffin, Z. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language, 56*, 521-536.

Arnold, J. E., Bennetto, L., & Diehl, J. J. (2009). Reference production in young speakers with and without autism: Effects of discourse status and processing constraints. *Cognition, 110*, 131-146.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition, 76*, B13-B26.

Arnold, J. E., & Tanenhaus, M. K. (2011). Disfluency effects in comprehension: How new information can become accessible. In E. Gibson & N. Perlmutter (Eds.), *The processing and acquisition of reference* (pp. 197-218). Cambridge, MA: MIT Press.

Arnold, J. E., Wasow, T., Asudeh, A., & Alrenga, P. (2004). Avoiding attachment ambiguities: The role of constituent ordering. *Journal of Memory and Language, 51*, 55-70.

Arts, J. P. M. (2004). *Overspecification in instructive texts.* PhD Dissertation, Tilburg University.

Audring, J. (2006). Pronominal Gender in Spoken Dutch. *Journal of Germanic Linguistics 18*, 85-116.

Audring, J. (2009). *Reinventing pronoun gender*. PhD Dissertation, Vrije Universiteit, Amsterdam.

Aylett, M., Turk, A. (2004). The Smooth Signal Redundancy Hypothesis: A Functional Explanation for Relationships between Redundancy, Prosodic Prominence, and Duration in Spontaneous Speech. *Language and Speech, 47*, 31-56.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Baddeley, A. (1986). *Working Memory*. Oxford: Clarendon Press.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 8). New York/London: Academic Press.

Bard, E. G., & Aylett, M. P. (2005). Referential Form, Word Duration, and Modeling the Listener in Spoken Dialogue. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to Studying World-Situated Language Use: Bridging the Language-as-Product and Language-as-Action Traditions* (pp. 173-191). Cambridge, MA: MIT Press.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the Intelligibility of Referring Expressions in Dialogue. *Journal of Memory and Language, 42*, 1-22.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Bates, D., Maechler, M., & Bolker, B. (2013). lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-2. URL: http://CRAN.R-project.org/package=lme4.

Bell, A., Jurafsky, D., Fosier-Lussier, E., Girand, C. G., Michelle, & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America, 113*(2), 1001-1024.

Belz, A., Kow, E., Viethen, J., & Gatt, A. (2010). Generating Referring Expressions in Context: The GREC Task Evaluation Challenges. In E. Krahmer & M. Theune (Eds.), *Empirical Methods in NLG, NLAI 5790* (pp. 294-327). Berlin/Heidelberg: Springer.

Van Bergen, G. (2011). *Who's first and what's next. Animacy and word order variation in Dutch language production.* PhD Dissertation, Radboud University Nijmegen.

Van Bergen, G., Stoop, W., Vogels, J., & De Hoop, H. (2011). Leve *hun*! Waarom hun nog steeds *hun* zeggen. *Nederlandse Taalkunde, 16*, 2-29.

Beun, R.-J., & Cremers, A. H. M. (1998). Object reference in a shared domain of conversation. *Pragmatics & Cognition, 6*(1/2), 121-152.

Bock, J. K. (1982). Toward a Cognitive Psychology of Syntax: Information Processing Contributions to Sentence Formulation. *Psychological Review, 89*(1), 1-47.

Bock, J. K., & Irwin, D. E. (1980). Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 467-484.

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition, 21*, 47-67.

Bock, J. K., Irwin, D. E., & Davidson, D. J. (2004). Putting First Things First. In J. M. Henderson & F. Ferreira (Eds.), *The Interface of Language, Vision and Action. Eye Movements and the Visual World* (pp. 249-278). New York/Hove: Psychology Press.

Bock, J. K., Loebell, H., & Morey, R. (1992). From Conceptual Roles to Structural Relations: Bridging the Syntactic Cleft. *Psychological Review, 99*(1), 150-171.

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, Syntax, and Semantics. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in Mind: Advances in the study of Language and Thought*. Cambridge, MA: MIT Press.

Bouma, G. J. (2008). *Starting a sentence in Dutch: a corpus study of subject- and object-fronting*. PhD Dissertation, Rijksuniversiteit Groningen.

Branigan, H. P., & Feleki, E. (1999). Conceptual Accessibility and Serial Order in Greek Speech Production. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society*. Vancouver, British Columbia.

Branigan, H. P., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua, 118*, 172-189.

Brennan, S. E. (1995). Centering Attention in Discourse. *Language and Cognitive Processes, 10*(2), 137-167.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(6), 1482-1493.

Brennan, S. E., & Hanna, J. E. (2009). Partner-Specific Adaptation in Dialog. *Topics in Cognitive Science, 1*, 274–291.

Brennan, S. E., Friedman, M. W., & Pollard, C. J. (1987). A Centering Approach to Pronoun *Proceedings of the 25th annual meeting on Association for Computational Linguistics* (pp. 155-162). Stroudsburg, PA: Association for Computational Linguistics.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the Dative Alternation. In G. Bouma, I. Kraemer & J. Zwarts (Eds.), *Cognitive Foundations of Interpretation* (pp. 69-94): Royal Netherlands Academy of Science.

Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language, 54*, 592-609.

Brown-Schmidt, S., Byron, D. K., & Tanenhaus, M. K. (2005). Beyond salience: Interpretation of personal and demonstrative pronouns. *Journal of Memory and Language, 53*, 292-313.

Brunetti, L. (2009). On the semantic and contextual factors that determine topic selection in Italian and Spanish. *The Linguistic Review, 26*(2-3), 261-289.

Burzio, L. (1998). Anaphora and Soft Constraints. In P. B. D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky (Eds.), *Is the Best Good Enough?: Optimality and Competition in Syntax*. Boston: MIT.

Callaway, C. B., & Lester, J. C. (2002). Pronominalization in generated discourse and dialogue. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics* (pp. 88-95). Association for Computational Linguistics.

Cardinaletti, A., Starke, M. (1996). Deficient pronouns: A view from Germanic. A study in the unified description of Germanic and Romance. In H. Thráinsson, S.D. Epstein, S. Peter (Eds.), *Studies in Comparative Germanic Syntax*. Dordrecht: Kluwer.

Carlson, G. N. (1977). A unified analysis of the English bare plural. *Linguistics and philosophy, 1*, 413-456.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25-56). New York: Academic Press.

Chafe, W. L. (1994). *Discourse, consciousness, and time: the flow and displacement of conscious experience in speaking and writing*. Chicago: The University of Chicago Press.

Chafe, W. L. (1996). Inferring Identifiability and Accessibility. In T. Fretheim & J. K. Gundel (Eds.), *Reference and Referent Accessibility* (pp. 37-46). Amsterdam/Philadelphia: John Benjamins.

Chiriacescu, S. I. (2011). Effects of Reference Form on Frequency of Mention and Rate of Pronominalization. In I. Hendrickx, S. L. Devi, A. Branco & R. Mitkov (Eds.), *Anaphora Processing and Applications, Lecture Notes in Artificial Intelligence* (pp. 132-143). Berlin/Heidelberg: Springer.

Christianson, K., & Ferreira, F. (2005). Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition, 98*, 105-135.

Clark, H. H. (1973). The Language-as-Fixed-Effect Fallacy: A Critique of Language Statistics in Psychological Research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335-359.

Clark, H. H. (1996). *Using Language*. Cambridge: University Press.

Clark, H. H., & Bangerter, A. (2004). Changing Ideas about Reference. In Noveck & D. Sperber (Eds.), *Experimental Pragmatics* (pp. 25-49). Basingstoke: Palgrave Macmillan.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3*(3), 472-517.

Clark, H. H., & Haviland, S. E. (1977). Comprehension and the Given-New contract. In R. O. Freedle (Ed.), *Discourse Production and Comprehension* (pp. 1-40). Norwood, N. J.: Ablex Publishing.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber & I. A. Sag (Eds.), *Elements of Discourse Understanding* (pp. 10-63). Cambridge: Cambridge University Press.

Clark, H. H., & Sengul, C. L. (1979). In search of referents for nouns and pronouns. *Memory and Cognition, 7*, 35-41.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1-39.

Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common Ground and the Understanding of Demonstrative Reference. *Journal of Verbal Learning and Verbal Behavior, 22*, 245-258.

Coco, M. I., & Keller, F. (2009). The impact of visual information on reference assignment in sentence production. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 274-279). Austin, TX: Cognitive Science Society.

Coco, M. I., & Keller, F. (2010). Sentence Production in Naturalistic Scenes with Referential Ambiguity. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 1070-1075). Austin, TX: Cognitive Science Society.

Comrie, B. (1989). *Language universals and linguistic typology* (2nd ed.). Chicago: University of Chicago Press.

Coppen, P.-A., Haeseryn, W., De Vriend, F. (2002). De Elektronische ANS, Version 1.2. Retrieved 18 February 2013 from http://ans.ruhosting.nl/e-ans/.

Corbett, G. (1991). *Gender*. University Press, Cambridge.

Croft, W. (1994). Voice: Beyond Control and Affectedness. In B. A. Fox & P. J. Hopper (Eds.), *Voice: Form and Function* (pp. 89-118). Amsterdam/Philadelphia: John Benjamins.

Dabrowska, E. (1998). How metaphor affects grammatical coding: the Saxon genitive in computer manuals. *English Language and Linguistics, 2*(1), 121-127.

Dahl, Ö. (2000). Animacy and the notion of semantic gender. In B. Unterbeck & M. Rissanen (Eds.), *Gender in Grammar and Cognition* (pp. 99-115). Berlin/New York: Mouton de Gruyter.

Dahl, Ö. (2008). Animacy and egophoricity: Grammar, ontology and phylogeny. *Lingua, 118*(2), 141-150.

Dahl, Ö., & Fraurud, K. (1996). Animacy in Grammar and Discourse. In T. Fretheim & J. Gundel (Eds.), *Reference and Referent Accessibility* (pp. 47-64). Amsterdam: John Benjamins.

Dale, R., & Reiter, E. (1995). Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science, 19*(2), 233-263.

Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener adaptation in language production. In D. Napoli & J. A. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore festschrift for Lila Gleitman* (pp. 105-129). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193-222.

Dowty, D. (1991). Thematic Proto-roles and Argument Selection. *Language, 67*(3), 547-619.

É. Kiss, K. (2002). *The Syntax of Hungarian*. Cambridge: University Press.

Engelhardt, P. E., Bailey, K. G. D., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language, 54*, 554-573.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective Taking as Egocentric Anchoring and Adjustment. *Journal of Personality and Social Psychology, 87*(3), 327-339.

Ferreira, V. S., & Yoshita, H. (2003). Given-New Ordering Effects on the Production of Scrambled Sentences in Japanese. *Journal of Psycholinguistic Research, 32*(6), 669-692.

Ferreira, V. S., Slevc, L. R., & Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition, 96*, 263–284.

Flores d'Arcais, G. B. (1975). Some perceptual determinants of sentence construction. In G. B. Flores d'Arcais (Ed.), *Studies in perception: Festschrift for Fabio Metelli* (pp. 344–373). Milan, Italy: Martello-Guinti.

Foraker, S., & McElree, B. (2007). The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language, 56*, 357-383.

Fraurud, K. (1996). Cognitive Ontology and NP Form. In T. Fretheim & J. K. Gundel (Eds.), *Reference and Referent Accessibility* (pp. 65-87). Amsterdam/Philadelphia: John Benjamins.

Fukumura, K., & Van Gompel, R. P. G. (2010). Choosing anaphoric expressions: Do people take into account likelihood of reference? *Journal of Memory and Language, 62*, 52-66.

Fukumura, K., & Van Gompel, R. P. G. (2011). The effect of animacy on the choice of referring expression. *Language and Cognitive Processes, 26*(10), 1472-1504.

Fukumura, K., & Van Gompel, R. P. G. (2012). Producing Pronouns and Definite Noun Phrases: Do Speakers Use the Addressee's Discourse Model? *Cognitive Science, 36*(7), 1289-1311.

Fukumura, K., Van Gompel, R. P. G., Harley, T., & Pickering, M. J. (2011). How does similarity-based interference affect the choice of referring expression? *Journal of Memory and Language, 65*, 331-344.

Fukumura, K., Van Gompel, R., & Pickering, M. J. (2010). The use of visual context during the production of referring expressions. *The Quarterly Journal of Experimental Psychology, 63*(9), 1700-1715.

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language, 62*(1), 35-51.

Gann, T. M., & Barr, D. J. (2012). Speaking from experience: Audience design as expert performance. *Language and Cognitive Processes*.

Geerts, G. (1966). *Genus en geslacht in de Gouden Eeuw. Een bijdrage tot de studie van de nominale klassifikatie en daarmee samenhangende adnominale flexievormen en pronominale verschijnselen in Hollands taalgebruik van de zeventiende eeuw*. Brussels: Belgisch Interuniversitair Centrum voor Neerlandistiek.

Gelman, R., Durgin, F., & Kaufman, L. (1995). Distinguishing between animates and inanimates: not by motion alone. In D. Sperber, D. Premack & A. J. Premack (Eds.), *Causal Cognition: A Multidisciplinary Debate*. Oxford: Clarendon Press.

Gernsbacher, M. A., & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language, 27*(6), 699-717.

Givón, T. (1976). Topic, Pronoun, and Grammatical Agreement. In C.N. Li (Ed.), *Subject and Topic* (pp. 149-188). New York/San Francisco/London: Academic Press.

Givón, T. (1983). *Topic Continuity in Discourse*. Amsterdam/Philadelphia: John Benjamins.

Givón, T. (1988). The pragmatics of word order: predictability, importance and attention. In M. Hammond, E. A. Moravcsik & J. R. Wirth (Eds.), *Studies in Syntactic Typology* (Vol. 17). Amsterdam/Philadelphia: John Benjamins.

Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the *give* and *take* between event apprehension and utterance formulation. *Journal of Memory and Language, 57*(4), 544-569.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names and the centering of attention in discourse. *Cognitive Science, 7*(3), 311-347.

Goudbeek, M. B., & Krahmer, E. J. (2011). Referring under load: Disentangling preference-based and alignment-based content selection processes in referring expression generation. In K. v. Deemter, A. Gatt, R. P. G. v. Gompel & E. J. Krahmer

(Eds.), *Proceedings of PRE-Cogsci: Bridging the gap between computational, empirical and theoretical approaches to reference*.

Greene, S. B., Gerrig, R. J., McKoon, G., & Ratcliff, R. (1994). Unheralded Pronouns and Management by Common Ground. *Journal of Memory and Language, 33*, 511-526.

Grice, H. P. (1975). Logic and Conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics. Vol. III: Speech acts* (pp. 41-58). New York: Academic Press.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11*, 274–279.

Grosz, B. J. (1977). *The representation and use of focus in a system for understanding dialogs*. Menlo Park, CA: Artificial Intelligence Center, SRI International.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse *Proceedings of the 21st Annual Meeting of the ACL* (pp. 44-50). Association of Computational Linguistics.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A Framework for Modelling the Local Coherence of Discourse. *Computational Linguistics, 21*(2), 203-225.

Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics, 12*(3), 175-204.

Gullberg, M. (2006). Handling Discourse: Gestures, Reference Tracking, and Communication Strategies in Early L2. *Language Learning, 56*(1), 155-196.

Gundel, J. K. (1988). Universals of topic-comment structure. In M. Hammond, E. A. Moravcsik & J. R. Wirth (Eds.), *Studies in Syntactic Typology* (Vol. 17). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of anaphoric expressions in discourse. *Language, 69*, 274-307.

Hawkins, J. A. (1994). *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

Heider, F., & Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American Journal of Psychology, 57*(2), 243-259.

Henderson, J. M., & Ferreira, F. (Eds.). (2004). *The Interface of Language, Vision, and Action. Eye Movements and the Visual World*. New York/Hove: Psychology Press.

Hendriks, P., Englert, C., Wubs, E., & Hoeks, J. (2008). Age Differences in Adults' Use of Referring Expressions. *Journal of Logic, Language and Information, 17*, 443-466.

Hendriks, P., Koster, C., & Hoeks, J. C. J. (2013). Referential choice across the lifespan: why children and elderly adults produce ambiguous pronouns. *Language and Cognitive Processes*.

Hoetjes, M., Koolen, R., Goudbeek, M., Krahmer, E., & Swerts, M. (2011). GREEBLES Greeble greeb: On Reduction in Speech and Gesture in Repeated References. In L. Carlson, C. Hoelscher & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3250-3255). Austin, TX: Cognitive Science Society.

Horton, W. S., & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. *Cognition, 96*, 127–142.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59*, 91-117.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford/New York: Oxford University Press.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434-446.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology, 61*, 23-62.

Jaeger, T. F. (2011). Post to HLP/Jaeger lab blog, 25 June 2011. http://hlplab.wordpress.com/2011/06/25/more-on-random-slopes

Jarvella, R. J., & Klein, W. (Eds.) (1982). *Speech, place, and action: Studies of deixis and related topics*. New York: Wiley.

Johnson-Laird, P. N. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Josefsson, G. (2010). "Disagreeing" pronominal reference in Swedish and the interplay between formal and semantic gender. *Lingua, 120*, 2095-2120.

Kaiser, E. (2010). Effects of Contrast on Referential Form: Investigating the Distinction Between Strong and Weak Pronouns. *Discourse Processes, 47*, 480-509.

Kaiser, E. (2011). Salience and contrast effects in reference resolution: The interpretation of Dutch pronouns and demonstratives. *Language and Cognitive Processes, 26*, 1587-1624.

Kaiser, E., Li, D. C.-H., Holsinger, E. (2011). Exploring the Lexical and Acoustic Consequences of Referential Predictability. In I. Hendricks, A. Branco, S. Lalitha Devi, R. Mitkov (Eds.), *Anaphora Processing and Applications, Lecture Notes in Artificial Intelligence* (pp. 171-183). Berlin/Heidelberg: Springer.

Kaiser, E., Trueswell, J. C. (2004). The referential properties of Dutch pronouns and demonstratives: Is salience enough?, *Proceedings of the Sinn und Bedeutung 8*, Arbeitspapier Nr. 1977. FB Sprachwissenschaft, Universität Konstanz.

Kaiser, E., Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes, 23*, 709-748.

Kantola, L., & Van Gompel, R. P. G. (2011). Does the addressee matter when choosing referring expressions? In K. v. Deemter, A. Gatt, R. P. G. v. Gompel & E. J. Krahmer (Eds.), *Proceedings of PRE-Cogsci: Bridging the gap between computational, empirical and theoretical approaches to reference*.

Karmiloff-Smith, A. (1981). The grammatical marking of thematic structure in the development of language production. In W. Deutsch (Ed.), *The Child's Construction of Language* (pp. 121–148). London: Academic Press.

Katz, J. J. (1972). *Semantic Theory*. New York: Harper & Row, Publishers.

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and Coreference Revisited. *Journal of Semantics, 25*(1), 1-44.

Kelleher, J., Costello, F., & Van Genabith, J. (2005). Dynamically structuring, updating and interrelating representations of visual and linguistic discourse context. *Artificial Intelligence, 167*, 62-102.

Kellogg, R. T., Olive, T., & Piolat, A. (2007). Verbal, visual, and spatial working memory in written language production. *Acta Psychologica, 124*, 382–397.

Kempen, G., & Harbusch, K. (2004). A corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In T. Pechmann & C. Habel (Eds.), *Multidisciplinary approaches to language production*. Berlin: Mouton De Gruyter.

Kempen, G., & Hoenkamp, E. (1987). An Incremental Procedural Grammar for Sentence Formulation. *Cognitive Science, 11*, 201-258.

Keysar, B., Barr, D. J., & Horton, W. S. (1998). The Egocentric Basis of Language Use: Insights from a Processing Approach. *Current Directions in Psychological Science, 7*(2), 46-50.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a Model of Text Comprehension and Production. *Psychological Review, 85*(5), 363-394.

Kirsner, R. S. (1979). Deixis in Discourse: An Exploratory Quantitative Study of the Modern Dutch Demonstrative Adjectives in Discourse and Syntax. *Syntax and Semantics, 12*, 355-375.

Koolen, R. M. F., Gatt, A., Goudbeek, M. B., & Krahmer, E. J. (2011). Factors causing overspecification in definite descriptions. *Journal of Pragmatics, 43*, 3231-3250.

Koolen, R. M. F., Goudbeek, M. B., & Krahmer, E. J. (2013). The Effect of Scene Variation on the Redundant Use of Color in Definite Reference. *Cognitive Science, 37*, 395-411.

Koolen, R. M. F., Krahmer, E. J., & Swerts, M. G. J. (2013). The impact of bottom-up and top-down saliency cues on reference production. In M. Knauff, M. Pauen, N. Sebanz & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 817-822). Austin, TX: Cognitive Science Society.

Krahmer, E. J., & Van Deemter, K. (2012). Computational Generation of Referring Expressions: A Survey. *Computational Linguistics, 38*(1), 173-218.

Krahmer, E. J., & Theune, M. (2002). Efficient Context-Sensitive Generation of Referring Expressions. In K. v. Deemter & R. Kibble (Eds.), *Information sharing: Reference and Presupposition in Language Generation and Interpretation* (pp. 223-264). Stanford, CA: CSLI Publications.

Kuno, S., & Kaburaki, E. (1977). Empathy and Syntax. *Linguistic Inquiry, 8*(4), 627-672.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Lambrecht, K. (1994). *Information Structure and Sentence Form. Topic, Focus, and the Mental Representations of Discourse Referents*. New York: Cambridge University Press.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge/London: The MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and Brain Sciences, 22*(1), 1-75.

Levinson, S. C. (1987). Pragmatics and the grammar of anaphora: A partial pragmatic reduction of binding and control phenomena. *Journal of Linguistics, 23*, 379-434.

MacWhinney, B. (1977). Starting Points. *Language, 53*(1), 152-168.

Maes, A. A. (1997). Referent Ontology and Centering in Discourse. *Journal of Semantics, 14*, 207-235.

Maes, A. A., & Noordman, L. G. M. (1995). Demonstrative nominal anaphors: a case of nonidentificational markedness. *Linguistics, 33*, 255-282.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2006). The effect of perceptual availability and prior discourse on young children's use of referring expressions. *Applied Psycholinguistics, 27*, 403-422.

Mazza, V., Turatto, M., & Umiltà, C. (2005). Foreground-background segmentation and attention: A change blindness study. *Psychological Research, 69*(3), 201-210.

McCoy, K., & Strube, M. (1999). Generating anaphoric expressions: Pronoun or definite description. In *Proceedings of ACL workshop on Discourse and Reference Structure* (pp. 63-71).

McDonald, J., Bock, J. K., & Kelly, M. H. (1993). Word and World Order: Semantic, Phonological and Metrical determinants of Serial Position. *Cognitive Psychology, 25*, 188-230.

Meyer, A. S., & Bock, J. K. (1999). Representations and Processes in the Production of Pronouns: Some Perspectives from Dutch. *Journal of Memory and Language, 41*, 281-301.

Mol, E. M. M., Krahmer, E. J., Maes, A. A., & Swerts, M. G. J. (2011). Seeing and Being Seen: The Effects on Gesture Production. *Journal of Computer-Mediated Communication, 17*(1), 77-100.

Montag, J. L., MacDonald, M. C. (2013). Visual Salience Modulates Structure Choice in Relative Clause Production. *Language and Speech*.

Morrow, D. G. (1985). Prominent Characters and Events Organize Narrative Understanding. *Journal of Memory and Language, 24*(3), 304-319.

Myachykov, A., Garrod, S. C., & Scheepers, C. (2009). Attention and syntax in sentence production: A critical review. *Discours, 4*. URL: http://discours.revues.org/index7594.html

New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *PNAS, 104*(42), 16598-16603.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When Peanuts Fall in Love: N400 Evidence for the Power of Discourse. *Journal of Cognitive Neuroscience, 18*(7), 1098-1111.

Nunberg, G. (1979). The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy, 3*(2), 143-184.

Osgood, C. E. (1971). Where do sentences come from? In D. D. Steinberg & L. A. Jakobovits (Eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics and psychology* (pp. 497-529). London: Cambridge University Press.

Osgood, C. E., & Bock, J. K. (1977). Salience and sentencing: some production principles. In S. Rosenberg (Ed.), *Sentence production: developments in research and theory* (pp. 89-145). Hillsdale, NJ: Erlbaum.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research, 42*(1), 107-123.

Pickering, M., J. & Garrod, S. C. (2004). Towards a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*, 169-226.

Pickering, M. J., & Garrod, S. C. (2013). How tightly are production and comprehension interwoven? *Frontiers in Psychology, 4*:238.

Piwek, P., Beun, R.-J., & Cremers, A. H. M. (2008). 'Proximal' and 'distal' in language and cognition: Evidence from deictic demonstratives in Dutch. *Journal of Pragmatics, 40*(4), 694-718.

Poesio, M., Stevenson, R., Di Eugenio, B., & Hitzeman, J. (2004). Centering: A Parametric Theory and Its Instantiations. *Computational Linguistics, 30*(3), 309-363.

Prat-Sala, M., & Branigan, H. P. (2000). Discourse constraints on syntactic processing in language production: a cross-linguistic study in English and Spanish. *Journal of Memory and Language, 42*, 168-182.

Pratt, J., Radulescu, P. V., Guo, R. M., & Abrams, R. A. (2010). It's Alive! Animate Motion Captures Visual Attention. *Psychological Science, 21*(11), 1724-1730.

Prince, E. F. (1981). Toward a Taxonomy of Given-New Information. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223-255). New York: Academic Press.

Reinhart, T. (1982). Pragmatics and Linguistics: An Analysis of Sentence Topics. *Philosophica, 27*(1), 53-94.

Van Rij, J. C. (2012). *Pronoun processing: Computational, behavioral, and psychophysiological studies in children and adults.* PhD Dissertation, University of Groningen.

Rohde, H. (2008). *Coherence-driven effects in sentence and discourse processing.* PhD Dissertation, UC San Diego.

Rosa, E. C., & Arnold, J. E. (2011). The Role of Attention in Choice of Referring Expression. In K. v. Deemter, A. Gatt, R. P. G. v. Gompel & E. J. Krahmer (Eds.), *Proceedings of PRE-Cogsci: Bridging the gap between computational, empirical and theoretical approaches to reference*.

Rosa, E. C., Finch, K. H., Bergeson, M., & Arnold, J. E. (2013). The effects of addressee attention on prosodic prominence. *Language and Cognitive Processes*.

Rosenbach, A. (2005). Animacy versus weight as determinants of grammatical variation in English. *Language, 81*, 613-644.

Rosenbach, A. (2008). Animacy and grammatical variation—Findings from English genitive variation. *Lingua, 118*, 151-171.

Sanford, A. J., & Garrod, S. C. (1981). *Understanding written language: Explorations of comprehension beyond the sentence*. Chichester/New York: Wiley.

Santos, N. S., David, N., Bente, G., & Vogeley, K. (2008). Parametric induction of animacy experience. *Consciousness and Cognition, 17*, 425-237.

de Saussure, F. (1959). *Course in General Linguistics* (W. Baskin, Trans.). New York: Philosophical Library (original work published 1916).

Schmitt, B. M., Meyer, A. S., & Levelt, W. J. M. (1999). Lexical access in the production of pronouns. *Cognition, 69*(3), 313-335.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4(8), 299-309.*

Schultz, J., Friston, K. J., O'Doherty, J., Wolpert, D. M., & Frith, C. D. (2005). Activation in Posterior Superior Temporal Sulcus Parallels Parameter Inducing the Percept of Animacy. *Neuron, 45*, 625-635.

Sedivy, J. C. (2003). Pragmatic Versus Form-Based Accounts of Referential Contrast: Evidence for Effects of Informativity Expectations. *Journal of Psycholinguistic Research, 32*(1), 3-23.

Van der Sijs, N. (2004). *Taal als mensenwerk: het ontstaan van het ABN*. SDU, Den Haag.

Sridhar, S. N. (1988). *Cognition and sentence production: a cross-linguistic study*. New York: Springer.

Stevenson, R. J. (2002). The Role of Salience in the Production of Referring Expressions. In K. v. Deemter & R. Kibble (Eds.), *Information sharing: Reference and Presupposition in Language Generation and Interpretation* (pp. 167–192). Stanford, CA: CSLI Publications.

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes, 94*, 473-592.

Strube, M., & Wolters, M. (2000). A probabilistic genre-independent model of pronominalization. In *Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference* (pp. 18-25). Association for Computational Linguistics.

De Swart, H. E., & De Hoop, H. (2000). Topic and Focus. In L. Cheng & R. Sybesma (Eds.), *The First Glot International State-of-the-Article Book* (pp. 105-130). Berlin/New York: Mouton de Gruyter.

Tomlin, R. S. (1986). *Basic word order: Functional principles*. London/Wolfeboro, N.H.: Croom Helm.

Tomlin, R. S. (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. In J. Nuyts & E. Pederson (Eds.), *Language and conceptualization* (pp. 162–189). New York: Cambridge University Press.

Tremoulet, P. D., & Feldman, J. (2000). Perception of animacy from the motion of a single object. *Perception*, *29*, 943-951.

Van Nice, K. Y., & Dietrich, R. (2003a). Animacy effects in language production: From mental model to formulator. In H. Härtl & H. Tappe (Eds.), *Mediating between Concepts and Grammar* (pp. 101-117). Berlin/New York: Mouton de Gruyter.

Van Nice, K. Y., & Dietrich, R. (2003b). Task sensitivity of animacy effects: evidence from German picture descriptions. *Linguistics, 41*(5), 825-849.

Veenker, T. (2003). WWStim. Een CGI-script voor het uitvoeren van eenvoudige experimenten via het world wide web, version 1.4.4. Universiteit Utrecht.

Van Vliet, S. M. K. (2008). *Proper Nouns and Pronouns. The production of referential expressions in narrative discourse*. PhD Dissertation, Utrecht: LOT.

De Vogelaer, G. (2006). Pronominaal genus bij 'Zuid-Nederlandse' taalverwervers: grammaticaal of semantisch systeem? In M. Hüning, U. Vogl, T. v. d. Wouden, A. Verhagen (Eds.), *Nederlands tussen Duits en Engels* (pp. 89-102). Leiden: Stichting Neerlandistiek Leiden.

De Vogelaer, G. (2009). Frequency effects in gender change: data from East and West Flemish dialects. In L. Aelbrecht, F. Brisard, P. Dendale, D. Jaspers, B. Le Bruyn, (Eds.), *Papers of the Linguistic Society of Belgium 3*. Brussels: Linguistic Society of Belgium.

De Vogelaer, G., & De Sutter, G. (2011). The geography of gender change: pronominal and adnominal gender in Flemish dialects of Dutch. *Language Sciences 33*, 192-205.

Vogels, J., Krahmer, E. J., & Maes, A. A. (2011). How visual saliency affects referent accessibility. In L. Carlson, C. Hoelscher & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3128-3133). Austin, TX: Cognitive Science Society.

Vogels, J., & Van Bergen, G. (2013). Where to place the subject in Dutch: The role of prototypical topic properties. *Corpus Linguistics and Linguistic Theory*.

Vonk, W., Hustinx, L. G. M. M., & Simons, W. H. G. (1992). The Use of Referential Expressions in Structuring Discourse. *Language and Cognitive Processes, 7*(3/4), 301-333.

Wardlow Lane, L., Groisman, M., & Ferreira, V. S. (2006). Don't talk about pink elephants! Speakers' control over leaking private information during language production. *Psychological Science, 17*(4), 273-277.

Watson, D. G., Arnold, J. E., Tanenhaus, M. K. (2008). Tic Tac TOE: Effects of predictability and importance on acoustic prominence in language production. *Cognition, 106*, 1548-1557.

Van der Wege, M. M. (2009). Lexical entrainment and lexical differentiation in reference phrase choice. *Journal of Memory and Language, 60*, 448-463.

Wiese, H. (2003). Semantics as a gateway to language. In H. Härtl & H. Tappe (Eds.), *Mediating between Concepts and Grammar* (pp. 197-222). Berlin/New York: Mouton de Gruyter.

Wubs, E., Hendriks, P., Hoeks, J., & Koster, C. (2009). Tell me a story! Children's capacity for topic shift *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 3)* (pp. 313-324). Somerville, MA: Cascadilla Press.

Yamamoto, M. (1999). *Animacy and Reference: A cognitive approach to corpus linguistics*. PhD Dissertation, Amsterdam/Philadelphia: John Benjamins.

# Appendices

## Appendix A: Experimental materials from Chapter 3

**Table A1.** Animate and inanimate lexical items used in Experiment 1.

| Animate | | Inanimate | |
|---|---|---|---|
| *schaatser* | 'skater' | *handtas* | 'handbag' |
| *prins* | 'prince' | *voetbal* | 'football' |
| *peuter* | 'toddler' | *schoen* | 'shoe' |
| *visser* | 'fisherman' | *fles* | 'bottle' |
| *padvinder* | 'boy scout' | *steen* | 'stone' |
| *koningin* | 'queen' | *sneeuwvlok* | 'snowflake' |
| *danseres* | '(female) dancer' | *lamp* | 'lamp' |
| *boef* | 'scoundrel' | *eierdoos* | 'egg carton' |

**Table A2.** Nonsense words used in Experiment 2.

*daptinder*
*knurp*
*kilper*
*vopper*
*pundimper*
*sloeiweurd*
*krielf*
*etalbuns*

**Appendix B: Experimental materials from Chapter 4**

(1) a. *Plotseling valt de klok tegen de inbreekster aan. Is ____ nu verraden?*
b. *Plotseling valt de inbreekster tegen de klok aan. Is ____ nu verraden?*
c. *Plotseling valt de klok tegen de inbreekster aan. Is ____ kapot?*
d. *Plotseling valt de inbreekster tegen de klok aan. Is ____ kapot?*
'Suddenly, the clock/burglaress falls against the burglaress/clock. Is ____ broken/now betrayed?'

(2) a. *Straks valt de vaas tegen de koopman aan. Is ____ onvoorzichtig?*
b. *Straks valt de koopman tegen de vaas aan. Is ____ onvoorzichtig?*
c. *Straks valt de vaas tegen de koopman aan. Is ____ duur?*
d. *Straks valt de koopman tegen de vaas aan. Is ____ duur?*
'Next thing we know, the vase/merchant will fall against the merchant/vase. Is ____ careless/expensive?'

(3) a. *Eindelijk past de steunkous de verpleegster toch. Heeft ____ nog pijn?*
b. *Eindelijk past de verpleegster de steunkous toch. Heeft ____ nog pijn?*
c. *Eindelijk past de steunkous de verpleegster toch. Zit ____ niet te strak?*
d. *Eindelijk past de verpleegster de steunkous toch. Zit ____ niet te strak?*
'Finally the support stocking/nurse fits the nurse/support stocking after all. Is ____ still in pain/not too tight?'

(4) a. *Misschien past de broek de zakenman niet. Is ____ te dik?*
b. *Misschien past de zakenman de broek niet. Is ____ te dik?*
c. *Misschien past de broek de zakenman niet. Is ____ te wijd?*
d. *Misschien past de zakenman de broek niet. Is ____ te wijd?*
'Perhaps the trousers/businessman won't fit the businessman/trousers. Is ____ too fat/too wide?'

(5) a. *'s Avonds warmt de kamer de waardin op. Had ____ het koud?*
b. *'s Avonds warmt de waardin de kamer op. Had ____ het koud?*
c. *'s Avonds warmt de kamer de waardin op. Is ____ wel goed geïsoleerd?*
d. *'s Avonds warmt de waardin de kamer op. Is ____ wel goed geïsoleerd?*
'At night the room/hostess warms up the hostess/room. Was ____ cold?/Is ____ well insulated?'

(6) a. *'s Middags verwarmt de soep de huisman snel. Is ____ tevreden?*
b. *'s Middags verwarmt de huisman de soep snel. Is ____ tevreden?*
c. *'s Middags verwarmt de soep de huisman snel. Smaakt ____ goed?*
d. *'s Middags verwarmt de huisman de soep snel. Smaakt ____ goed?*
'At noon the soup/househusband warms the househusband/soup quickly. Is ____ pleased?/Does ____ taste good?'

(7) a. *Even later passeert de kar de heks opnieuw. Is ____ verdwaald?*
b. *Even later passeert de heks de kar opnieuw. Is ____ verdwaald?*
c. *Even later passeert de kar de heks opnieuw. Is ____ zwaar beladen?*
d. *Even later passeert de heks de kar opnieuw. Is ____ zwaar beladen?*
'After a while, the cart/witch passes the witch/cart again. Is ____ lost/heavily-laden?'

(8) a. *Met gemak zet de hijskraan de reus overeind. Is ____ gespierd?*
b. *Met gemak zet de reus de hijskraan overeind. Is ____ gespierd?*
c. *Met gemak zet de hijskraan de reus overeind. Is ____ van stevig metaal gemaakt?*
d. *Met gemak zet de reus de hijskraan overeind. Is ____ van stevig metaal gemaakt?*
'With ease the hoisting crane/giant stands up the giant/hoisting crane. Is ____ muscular/made of solid metal?'

(9) a. *Vervelend genoeg valt de kast tegen de secretaresse aan. Is ____ gewond?*
b. *Vervelend genoeg valt de secretaresse tegen de kast aan. Is ____ gewond?*
c. *Vervelend genoeg valt de kast tegen de secretaresse aan. Heeft ____ scherpe randjes?*
d. *Vervelend genoeg valt de secretaresse tegen de kast aan. Heeft ____ scherpe randjes?*
'Unfortunately the cupboard/secretary falls against the secretary/cupboard. Is ____ hurt?/Does ____ have sharp edges?'

(10) a. *Met slecht weer beschermt de tent de cameraman goed. Moet ____ de hele dag filmen?*
b. *Met slecht weer beschermt de cameraman de tent goed. Moet ____ de hele dag filmen?*
c. *Met slecht weer beschermt de tent de cameraman goed. Is ____ gemakkelijk verplaatsbaar?*
d. *Met slecht weer beschermt de cameraman de tent goed. Is ____ gemakkelijk verplaatsbaar?*
'In bad weather, the tent/cameraman shelters the cameraman/tent well. Does ____ have to shoot all day?/Can ____ be moved easily?'

(11) a. *Weer raakt de stemvork de muzieklerares aan. Hoort ____ een zuivere toon?*
b. *Weer raakt de muzieklerares de stemvork aan. Hoort ____ een zuivere toon?*
c. *Weer raakt de stemvork de muzieklerares aan. Geeft ____ een zuivere toon?*
d. *Weer raakt de muzieklerares de stemvork aan. Geeft ____ een zuivere toon?*
'Again, the tuning fork/music teacher hits the music teacher/tuning fork. Does ____ hear a clear tone/give a clear tone?'

(12) a. *Gelukkig raakt de kom de grootvader net niet. Is ____ geschrokken?*
b. *Gelukkig raakt de grootvader de kom net niet. Is ____ geschrokken?*
c. *Gelukkig raakt de kom de grootvader net niet. Is ____ gebarsten?*
d. *Gelukkig raakt de grootvader de kom net niet. Is ____ gebarsten?*
'Fortunately, the bowl/grandfather just misses the grandfather/bowl. Is ____ startled/burst?'

(13) a. *Schokkend genoeg vergiftigt de naald de zuster met arseen. Heeft ____ vijanden?*
b. *Schokkend genoeg vergiftigt de zuster de naald met arseen. Heeft ____ vijanden?*
c. *Schokkend genoeg vergiftigt de naald de zuster met arseen. Is ____ erg scherp?*
d. *Schokkend genoeg vergiftigt de zuster de naald met arseen. Is ____ erg scherp?*
'Shockingly, the needle/nurse poisons the nurse/needle with arsenic acid. Does ____ have enemies?/Is ____ very sharp?'

(14) a. *Vooralsnog houdt de tafel de barman maar net. Doet ____ wel voorzichtig?*
b. *Vooralsnog houdt de barman de tafel maar net. Doet ____ wel voorzichtig?*
c. *Vooralsnog houdt de tafel de barman maar net. Heeft ____ stevige houten poten?*
d. *Vooralsnog houdt de barman de tafel maar net. Heeft ____ stevige houten poten?*
'As yet, the table/barkeeper is only just holding the barkeeper/table. Is ____ being careful?/Does ____ have sturdy wooden legs?'

(15) a. *Ineens botst de slee tegen de prinses aan. Heeft ____ zich pijn gedaan?*
b. *Ineens botst de prinses tegen de slee aan. Heeft ____ zich pijn gedaan?*
c. *Ineens botst de slee tegen de prinses aan. Is ____ stuk?*
d. *Ineens botst de prinses tegen de slee aan. Is ____ stuk?*
'Suddenly the sleigh/princess bumps into the princess/sleigh. Did ____ hurt herself?/Is ____ broken?'

(16) a. *Eigenlijk omschrijft de bladzijde de koning onjuist. Heeft ____ bewust de feiten verdraaid?*
b. *Eigenlijk omschrijft de koning de bladzijde onjuist. Heeft ____ bewust de feiten verdraaid?*
c. *Eigenlijk omschrijft de bladzijde de koning onjuist. Bevat ____ gevoelige informatie?*
d. *Eigenlijk omschrijft de koning de bladzijde onjuist. Bevat ____ gevoelige informatie?*
'In fact, the page/king describes the king/page inaccurately. Has ____ consciously distorted the facts?/Does ____ contain sensitive information?'

# Summary

An essential part of our communication through language involves referring to things: objects, people, locations, abstract concepts, and so on. This dissertation is concerned with two choices a speaker makes when referring: the choice of *what to refer to first in an utterance* and the choice of *what type of expression to use to refer to it* (e.g., 'the girl with the big earrings' or 'she'). The aim of this dissertation is to get more insight into the role the accessibility (or ease of retrieval) of mental representations plays in both of these referential choices in language production.

**Chapter 1** provides a theoretical background on how the two referential choices have been treated in previous research. Traditionally, effects of accessibility on the choice of referent for first mention have been ascribed to the amount of attention the *speaker* herself has allocated to a referent, which may be influenced by linguistic, perceptual and conceptual factors. By contrast, effects of accessibility on the choice of referring expression have been explained as the speaker's effort to select expressions in accordance with the information she has in common ground with her *addressee*. What is part of common ground is assumed to be largely determined by the linguistic context. Hence, structural properties of the referent in the directly preceding discourse, such as grammatical function and topichood, have been taken to make the greatest contribution to accessibility. This dissertation contains four experimental studies, all conducted in Dutch, in which it was investigated whether some of the non-linguistic and speaker-internal factors that have been found to influence the choice of referent for first mention also play a role in the choice of referring expression. If so, the question is how they interact with linguistic factors.

The question addressed in **Chapter 2** is how different types of salience interact in influencing referential choices. Salient entities are assumed to be more accessible in memory, which makes them more likely to be referred to first and to be referred to with an attenuated expression, such as a pronoun. The study presented in this chapter investigates the interaction between the salience of a referent in the perceptual context (visual foregrounding) and its salience in the linguistic context (grammatical function). More specifically, it is asked whether a non-linguistic factor such as visual salience can

affect the choice of referent as well as the choice of referring expression in the presence of a linguistic context.

Two spoken story completion experiments were conducted, in which speakers provided continuations of linguistic contexts based on picture pairs showing two characters. The visual salience of the characters was manipulated by placing one of them in the foreground and the other one in the background. Linguistic salience was manipulated by varying the grammatical function of the referents in the linguistic context. To reduce the chances that linguistic salience becomes so strong that it overrules any effects of visual salience, linguistic salience was moderated in Experiment 1 by inducing a topic shift in the discourse context, in which one character was the subject of the first context sentence and the other one the subject of the second context sentence. In Experiment 2, contexts in which linguistic salience was unclear were compared to contexts in which one of the characters was highly linguistically salient.

The results show that visually salient referents are more likely to be referred to first, independently of linguistic salience. However, they are not more likely to be pronominalized. By contrast, linguistically salient referents (subjects) are more likely to be referred to with pronouns in a subsequent utterance, but they are not more likely to be referred to first. Still, referents that have been introduced as the discourse topic (what the story is about) *are* more likely candidates for first mention. These findings are taken as evidence that the choice of what to refer to first is mainly driven by the speaker's global representation of what is most important to talk about, which may be influenced by high-level factors such as protagonisthood as well as by low-level visual salience. Conversely, the local accessibility of individual referents is not affected by this, being primarily driven by structural properties of the linguistic context. This is compatible with a view of language production in which utterance planning is influenced by conceptual and discourse factors rather than by low-level perceptual factors directly.

**Chapter 3** is concerned with the inherent salience of referents. More specifically, it investigates effects of animacy on the choice of referent and the choice of referring expression. Several earlier studies have provided evidence that animacy influences referential choices. On the one hand, it has been shown that animate referents are more likely to be mentioned as subjects than inanimate referents. On the other hand, animate referents have been found to be more frequently pronominalized than inanimate referents. These effects have mainly been analyzed as effects of conceptual

accessibility, i.e., animate entities have a higher activation in memory than inanimate entities. In this chapter, the question is raised whether these effects are driven only by lexical concepts, such that referents described by animate lexical items (e.g., 'toddler') are more accessible than referents described by inanimate lexical items (e.g., 'shoe'), or whether they can also be influenced by context-derived conceptualizations, such that referents that are perceived as animate in a particular context are more accessible than referents that are not.

Two animation-retelling experiments are presented in which speakers watched simple animations and retold them to an addressee. In Experiment 1, the target referent in the animations moved either in an animate way or in an inanimate way (perceptual animacy). In addition, it was given a lexical label that could either be animate or inanimate (lexical animacy). In Experiment 2, lexical labels consisted of nonsense words, to exclude influences of lexical animacy on perceptual animacy. If the effects of animacy are context-dependent, entities that are perceived as animate should yield more subject references and more pronouns than entities that are perceived as inanimate, irrespective of their lexical animacy. If the effects are tied to lexical concepts, entities described with animate lexical items should be mentioned as the subject and pronominalized more frequently than entities described with inanimate lexical items, irrespective of their perceptual animacy.

The results of Experiment 1 show that the two factors affect referential choices differently. Whereas entities that are perceived as animate are more likely to be pronominalized, they are not more likely to be mentioned as the subject. Conversely, entities with animate names (but not necessarily perceived as animate) are more likely to be mentioned as the subject, but not more likely to be pronominalized. These findings show that conceptualizations based on the perceptual context can influence the choice of referring expression, and that these may overrule lexical semantics. However, this effect should be dissociated from animacy effects on likelihood of next mention (as the subject), in which mainly lexical animacy seems to play a role. This dissociation may be partly due to the time course of linguistic processing in interaction with task-specific dependencies. For example, because lexical labels were presented before the start of the animations, the activation of lexical items while producing an utterance may have overruled conceptual accessibility in the choice of referent. The choice of referring expression might have been less affected by this.

**Chapter 4** zooms in further on effects of animacy on the choice of referring expression in Dutch. As noted in Chapter 3, the finding that animate entities tend to be referred to with more attenuated expressions than inanimate entities has been explained by their higher conceptual accessibility in memory. However, two previously untested claims made for Dutch suggest that the situation may be more complex. Firstly, it has been stated that full pronouns can only refer to animate entities, while reduced pronouns can also refer to inanimate entities. This seems to be inconsistent with the accessibility account, in which references to more accessible entities should be more reduced. Secondly, inanimate entities may also be pronominalized less to avoid gender-marked expressions when grammatical gender is unclear. The aim of this chapter is twofold: On the one hand, it examines the predictions of theories of accessibility on a more fine-grained scale of referential forms, by looking at the distribution of full and reduced pronouns. On the other hand, it investigates whether (lexically) animate entities are more likely to be pronominalized in Dutch, and to what degree this can be explained by a gender avoidance strategy rather than by accessibility.

Using a sentence completion task, reference production of Dutch speakers from the Netherlands that lack grammatical intuitions about masculine and feminine nominal gender was compared to Dutch speakers from Belgium that largely retain these intuitions. Speakers orally completed sentences by filling a gap with a referring expression that had either an animate or an inanimate antecedent. If gender avoidance plays a role in the choice to pronominalize, Netherlandic Dutch speakers are predicted to show a larger effect of animacy (i.e., more full noun phrases for inanimate antecedents) than Belgian Dutch speakers, who do not need to avoid gender-marked expressions. In addition, if references to more accessible entities are more reduced, reduced pronouns are predicted to be more frequent for animate than for inanimate antecedents.

The results showed a similar effect of animacy on pronominalization for the two groups, suggesting that gender avoidance cannot explain the animacy effect. Hence, this effect is likely to be mainly driven by conceptual accessibility. At the same time, reduced pronouns were more frequent for inanimate than for animate antecedents. This distribution cannot be explained by an accessibility account. Instead, the choice between a full and a reduced form may be affected by the importance of information for the speaker or the addressee. These results support the view that choices for different types of referring expressions are driven by different factors.

The aim of **Chapter 5** is to investigate whether the choice of referring expression is influenced by speaker-internal constraints. In traditional theories of reference, the general view is that speakers choose referring expressions such that their addressees can pick out the correct referent. More recent findings suggest that the choice of referring expression can also be influenced by speaker-internal factors, such as cognitive load. However, these findings do not always show a consistent pattern. Some researchers have argued that an increased cognitive load affects the amount of attention a speaker can allocate to a referent, which leads to a lower accessibility of the referent in the speaker's own memory. This results in more specific expressions. By contrast, others have argued that speakers with an increased cognitive load are less able to calculate accessibility for their addressees, which makes them more egocentric in their references. This might lead to an increase in the use of pronouns, because these expressions are assumed to be easiest to produce.

In two experiments, it was explored whether and how an increased cognitive load for speakers influences their choice between pronouns and more specific expressions, such as full noun phrases. Speakers produced story continuations to addressees, which required them to refer to characters in a picture that were either salient or non-salient in the discourse. To be able to verify whether speakers refer from the perspective of their addressee or from their own perspective, referents that were salient for the speaker were non-salient for the addressee, and vice versa, in Experiment 1. This setup was compared to a situation in which all discourse information was shared in Experiment 2. Cognitive load was manipulated by the presence or absence of a secondary verbal memory task: In one half of the experiment, speakers were presented with either the word BAL or the word DAL before and after each story, and had to indicate whether they had seen the same word ord not. In the other half of the experiment, there was no secondary task.

The effect of the increased cognitive load was that pronouns were used more often, at least when speakers referred to referents that were less salient for themselves. This was the case both when the referent's salience differed between the speaker's and the addressee's perspectives and when all discourse information was shared. This result suggests that accessibility is not identical to speaker attention. At the same time, speakers did not seem to take into account the referent's accessibility for the addressee either. It is proposed that speakers normally use their own discourse model as a proxy for their addressee's. However, when they experience an increased cognitive load,

speakers have more difficulties taking into account that referents that are less salient in this discourse model should be referred to with more specific expressions. This results in the use of expressions that are more economical for themselves.

Finally, **Chapter 6** relates the main findings in this dissertation to existing theories of reference production. While previous research on the production of referring expressions in discourse has mainly focused on linguistic and addressee-oriented cues for accessibility, the research in this dissertation provides evidence for the influence of both a non-linguistic factor (perceptual animacy) and a speaker-internal factor (cognitive load) on the choice of referring expression. These factors also interact with linguistic cues. At the same time, the studies presented in this dissertation make it clear that there are indeed differences in the effects of accessibility on different referential choices. Notably, factors that may influence the global conceptualization of a discourse, such as visual salience and protagonisthood, mainly affect the choice of referent for first mention, while local discourse factors such as grammatical function and topichood mainly affect the choice of referring expression. In addition, animacy seems to have quite disparate effects on different referential choices. These findings have consequences for our understanding of the notion of accessibility, since it suggests that there are multiple types or levels of accessibility, to which different language production processes can be sensitive in different ways. To get a better picture of referential choices in language production, the results of the present research must be incorporated into both theoretical and computational models of reference.

# List of publications

**Journal publications**

Vogels, J., Krahmer, E. J., & Maes, A. A. (submitted). How cognitive load influences speakers' choice of referring expressions.

Vogels, J., Maes, A. A., & Krahmer, E. J. (in press). Choosing referring expressions in Belgian and Netherlandic Dutch: effects of animacy. *Lingua*.

Vogels, J. & Van Bergen, G. (2013). Where to place inaccessible subjects in Dutch: The role of definiteness and animacy. *Corpus Linguistics and Linguistic Theory*.

Vogels, J., Krahmer, E. J., & Maes, A. A. (2013). When a stone tries to climb up a slope: the interplay between lexical and perceptual animacy in referential choices. *Frontiers in Psychology, 4*:154.

Vogels, J., Krahmer, E. J., & Maes, A. A. (2013). Who is where referred to how, and why? The influence of visual saliency on referent accessibility in spoken language production. *Language and Cognitive Processes, 28*(9), 1323-1349.

Van Bergen, G., Stoop, W., Vogels, J., & De Hoop, H. (2011). Leve *hun*! Waarom hun nog steeds *hun* zeggen. *Nederlandse Taalkunde, 16*(1), 2-29.

Vogels, J. & Lamers, M. J. A. (2008). The placement of bare plural subjects in Dutch. *Linguistics in the Netherlands, 25*, 169-80.

**Papers in conference proceedings (peer-reviewed)**

Vogels, J., Krahmer, E. J., & Maes, A. A. (2013). Cognitive load does not decrease pronoun use when speaker's and addressee's perspectives are dissociated. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 3681-3686). Austin, TX: Cognitive Science Society.

Vogels, J., Krahmer, E. J., & Maes, A. A. (2013). Effects of cognitive load on the choice of referential form. In A. Gatt, R. P. G. v. Gompel, E. G. Bard, E. J. Krahmer, & K. v. Deemter (Eds.), *Proceedings of PRE-CogSci: Bridging the gap between cognitive and computational approaches to reference, 31 July 2013, Berlin, Germany*.

Vogels, J., Krahmer, E. J., & Maes, A. A. (2011). How visual saliency affects referent accessibility. In L. Carlson, C. Hoelscher, & T.F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3128-3133). Austin, TX: Cognitive Science Society.

**Abstracts of conference presentations (peer-reviewed)**

Vogels, J., Krahmer, E. J., & Maes, A. A. (2012). When a stone tries to climb up a slope. The influence of perceived and linguistically induced animacy on reference. Poster presented at the *25th Annual CUNY Conference on Human Sentence Processing*. New York, NY.

Vogels, J., Krahmer, E. J., & Maes, A. A. (2010). The effect of visual salience on the production of referring expressions. Poster presented at *AMLaP 2010*, September 6-8, University of York.

# TiCC Ph.D. series

1. Pashiera Barkhuysen. *Audiovisual Prosody in Interaction.* Promotores: M.G.J. Swerts, E.J. Krahmer. Tilburg, 3 October 2008.
2. Ben Torben-Nielsen. *Dendritic Morphology: Function Shapes Structure.* Promotores: H.J. van den Herik, E.O. Postma. Co-promotor: K.P. Tuyls. Tilburg, 3 December 2008.
3. Hans Stol. *A Framework for Evidence-based Policy Making Using IT.* Promotor: H.J. van den Herik. Tilburg, 21 January 2009.
4. Jeroen Geertzen. *Dialogue Act Recognition and Prediction.* Promotor: H. Bunt. Co-promotor: J.M.B. Terken. Tilburg, 11 February 2009.
5. Sander Canisius. *Structured Prediction for Natural Language Processing.* Promotores: A.P.J. van den Bosch, W. Daelemans. Tilburg, 13 February 2009.
6. Fritz Reul. *New Architectures in Computer Chess.* Promotor: H.J. van den Herik. Co-promotor: J.W.H.M. Uiterwijk. Tilburg, 17 June 2009.
7. Laurens van der Maaten. *Feature Extraction from Visual Data.* Promotores: E.O. Postma, H.J. van den Herik. Co-promotor: A.G. Lange. Tilburg, 23 June 2009 (cum laude).
8. Stephan Raaijmakers. *Multinomial Language Learning.* Promotores: W. Daelemans, A.P.J. van den Bosch. Tilburg, 1 December 2009.
9. Igor Berezhnoy. *Digital Analysis of Paintings.* Promotores: E.O. Postma, H.J. van den Herik. Tilburg, 7 December 2009.
10. Toine Bogers. *Recommender Systems for Social Bookmarking.* Promotor: A.P.J. van den Bosch. Tilburg, 8 December 2009.
11. Sander Bakkes. *Rapid Adaptation of Video Game AI.* Promotor: H.J. van den Herik. Co-promotor: P. Spronck. Tilburg, 3 March 2010.
12. Maria Mos. *Complex Lexical Items.* Promotor: A.P.J. van den Bosch. Co-promotores: A. Vermeer, A. Backus. Tilburg, 12 May 2010 (in collaboration with the Department of Language and Culture Studies).
13. Marieke van Erp. *Accessing Natural History. Discoveries in data cleaning, structuring, and retrieval*. Promotor: A.P.J. van den Bosch. Co-promotor: P.K. Lendvai. Tilburg, 30 June 2010.

14. Edwin Commandeur. *Implicit Causality and Implicit Consequentiality in Language Comprehension*. Promotores: L.G.M. Noordman, W. Vonk. Co-promotor: R. Cozijn. Tilburg, 30 June 2010.
15. Bart Bogaert. *Cloud Content Contention*. Promotores: H.J. van den Herik, E.O. Postma. Tilburg, 30 March 2011.
16. Xiaoyu Mao. *Airport under Control.* Promotores: H.J. van den Herik, E.O. Postma. Co-promotores: N. Roos, A. Salden. Tilburg, 25 May 2011.
17. Olga Petukhova. *Multidimensional Dialogue Modelling*. Promotor: H. Bunt. Tilburg, 1 September 2011.
18. Lisette Mol. *Language in the Hands.* Promotores: E.J. Krahmer, A.A. Maes, M.G.J. Swerts. Tilburg, 7 November 2011 (cum laude).
19. Herman Stehouwer. *Statistical Language Models for Alternative Sequence Selection*. Promotores: A.P.J. van den Bosch, H.J. van den Herik. Co-promotor: M.M. van Zaanen. Tilburg, 7 December 2011.
20. Terry Kakeeto-Aelen. *Relationship Marketing for SMEs in Uganda*. Promotores: J. Chr. van Dalen, H.J. van den Herik. Co-promotor: B.A. Van de Walle. Tilburg, 1 February 2012.
21. Suleman Shahid. *Fun & Face: Exploring non-verbal expressions of emotion during playful interactions*. Promotores: E.J. Krahmer, M.G.J. Swerts. Tilburg, 25 May 2012.
22. Thijs Vis. *Intelligence, Politie en Veiligheidsdienst: Verenigbare Grootheden?* Promotores: T.A. de Roos, H.J. van den Herik, A.C.M. Spapens. Tilburg, 6 June 2012 (in collaboration with the Tilburg School of Law).
23. Nancy Pascall. *Engendering Technology Empowering Women.* Promotores: H.J. van den Herik, M. Diocaretz. Tilburg, 19 November 2012.
24. Agus Gunawan. *Information Access for SMEs in Indonesia*. Promotor: H.J. van den Herik. Co-promotores: M. Wahdan, B.A. Van de Walle. Tilburg, 19 December 2012.
25. Giel van Lankveld. *Quantifying Individual Player Differences.* Promotores: H.J. van den Herik, A.R. Arntz. Co-promotor: P. Spronck. Tilburg, 27 February 2013.
26. Sander Wubben. *Text-to-text Generation Using Monolingual Machine Translation.* Promotores: E.J. Krahmer, A.P.J. van den Bosch, H. Bunt. Tilburg, 5 June 2013.
27. Jeroen Janssens. *Outlier Selection and One-Class Classification*. Promotores: E.O. Postma, H.J. van den Herik. Tilburg, 11 June 2013.

28. Martijn Balsters. *Expression and Perception of Emotions: The Case of Depression, Sadness and Fear*. Promotores: E.J. Krahmer, M.G.J. Swerts, A.J.J.M. Vingerhoets. Tilburg, 25 June 2013.
29. Lisanne van Weelden. *Metaphor in Good Shape.* Promotor: A.A. Maes. Co-promotor: J. Schilperoord. Tilburg, 28 June 2013.
30. Ruud Koolen. *Need I say More? On Overspecification in Definite Reference*. Promotores: E.J. Krahmer, M.G.J. Swerts. Tilburg, 20 September 2013.
31. J. Douglas Mastin. *Exploring Infant Engagement. Language Socialization and Vocabulary. Development: A Study of Rural and Urban Communities in Mozambique.* Promotor: A.A. Maes. Co-promotor: Dr. P.A. Vogt. Tilburg, 11 October 2013.
32. Philip C. Jackson, Jr. *Toward Human-Level Artificial Intelligence – Representation and Computation of Meaning in Natural Language*. Promotores: H.C. Bunt, W.P.M. Daelemans. Tilburg, 22 April 2014.
33. Jorrig Vogels. *Referential choices in language production: The role of accessibility.* Promotores: A.A. Maes, E.J. Krahmer. Tilburg, 23 April 2014.